



Direct versus Communications-Based Prohibitions on Price Fixing

Citation

Louis Kaplow, Direct versus Communications-Based Prohibitions on Price Fixing, 3 J. Legal Analysis 449 (2011).

Published Version

<http://jla.oxfordjournals.org/content/3/2/449.full.pdf+html>

Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:30011624>

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

DIRECT VERSUS COMMUNICATIONS-BASED PROHIBITIONS ON PRICE FIXING

*Louis Kaplow**

ABSTRACT

This article compares two policies toward coordinated oligopolistic price elevation. Most commentators endorse the view that the law should (and does) prohibit only those price elevations produced by certain sorts of interfirm communications, such as secret price negotiations. In contrast, little attention has been devoted to a more direct approach that encompasses all coordinated price elevations that can be detected and sanctioned effectively. It is demonstrated that the conventional formulation rests on numerous misconceptions, involves complex and costly detection if its logical implications are taken seriously, and tends to target cases with relatively low deterrence benefits and high chilling costs in contrast to those targeted under the direct approach.

1. INTRODUCTION

The rule against price fixing is the least controversial prohibition in competition law throughout the world. There is, however, far less consensus than meets the eye on what constitutes price fixing and on how legal regimes should identify and remedy it. Moreover, prevailing views and existing doctrine are not grounded in systematic economic analysis of the problem. Instead, acting in a surprisingly formalistic manner, commentators—and, to a degree, government agencies and courts—have tended to focus on penalizing certain sorts of

* Finn M.W. Caspersen and Household International Professor of Law and Economics, Harvard Law School, and Research Associate, National Bureau of Economic Research. Address: Harvard Law School, HA322, Cambridge, MA 02138, USA. E-mail: meskridge@law.harvard.edu. I am grateful to Jonathan Baker, Aaron Edlin, Oliver Hart, Carl Shapiro, Steven Shavell, Kathryn Spier, the referees, and participants in workshops at Chicago, Columbia, Georgetown, Harvard, Michigan, NYU, Virginia, and Yale for helpful comments; Ariel Fox, Jesse Gurman, Lauren Merrell, Balaji Narain, Michael Sabin, and Dorothy Shapiro for research assistance; and Harvard Law School's John M. Olin Center for Law, Economics, and Business, the Leeds Research Fund, and the Kauffman Foundation for financial support. This article is part of a three-part project, the others of which are "On the Meaning of Horizontal Agreements in Competition Law" (Kaplow 2011c) and "An Economic Approach to Price Fixing" (Kaplow 2011a), and will ultimately be part of a book, tentatively titled "Competition Policy toward Price Fixing." Seeds for some of the ideas in this article appear in Areeda and Kaplow (1988) (and the Teacher's Manual) and Kaplow and Shapiro (2007).

interfirm communications that facilitate coordinated oligopolistic price elevation rather than on whether such price increases have occurred.¹ This article systematically compares this conventionally favored version of the prohibition to a more direct approach aimed at socially harmful coordinated price elevation that can be detected and sanctioned effectively.²

A priori, one would expect a direct approach to a problem to be superior to an indirect, circumscribed one. Analysts, enforcers, and adjudicators usually do best by asking the right question—the one of direct social concern—rather than by attempting to answer a substantially different one. Sometimes indirect approaches turn out to be superior, but this can be ascertained only after sustained analysis that articulates the competing methods and explicitly assesses their differences. It is therefore striking that many of the topics investigated here have received limited prior attention.

There is another respect in which the current state of discourse is puzzling. The proffered method involves the direct application of economic understandings to the problem at hand. In recent decades, courts and commentators have increasingly embraced the view that competition law should be grounded in economic substance rather than formalistic distinctions. The U.S. Supreme Court's decisions in *Sylvania*, *Matsushita*, and *Leegin* come to mind,³ and the European Union has been moving in a similar direction even if less strongly. Moreover, Richard Posner's (1969) article and subsequent book (2001, first edition 1976) famously endorsed a more economically based strategy for dealing with oligopoly pricing.⁴ Nevertheless, subsequent commentary, even when

-
- 1 This sort of position is prominently advanced by Turner (1962) and Areeda and Hovenkamp (2003). For an assessment of attempts to make sense of the law's concept of horizontal agreement from a number of perspectives, see Kaplow (2011c).
 - 2 Attention is confined to coordination on price. Nonprice coordination has many similarities but also important differences (for example, nonprice coordination is often more difficult to accomplish and detection of some types of coordinated behavior using market-based evidence may well be more challenging).
 - 3 *Continental T.V., Inc. v. GTE Sylvania Inc.*, 433 U.S. 36, 58–59 (1977); *Matsushita Elec. Indus. Co., Ltd. v. Zenith Radio Corp.*, 475 U.S. 574, 587, 596 (1986); *Leegin Creative Leather Products, Inc. v. PSKS, Inc.*, 551 U.S. 877, 887–889 (2007).
 - 4 If the economic evidence presented in a case warrants an inference of collusive pricing, there is neither legal nor practical justification for requiring evidence that will support the further inference that the collusion was explicit rather than tacit. From an economic standpoint it is a detail whether the collusive pricing scheme was organized and implemented in such a way as to generate evidence of actual communications. Posner (2001, 94).

citing his writings, does not significantly engage their content.⁵ As will be seen, the present analysis goes well beyond Posner's and differs in many respects, which can hardly be surprising in light of the decades of theoretical and empirical work in industrial organization in the interim; yet the spirit of his argument as well as some of his key ideas provide important illumination.

Part 2 presents the direct approach, beginning in section 2.1 with an examination of the social problem. From the outset, it is notable that none of the pertinent theory directly distinguishes between successful oligopolistic price elevation due merely to recognized interdependence (firms refrain from price cutting because of an expectation of retaliation derived from a shared appreciation of their circumstances) and that resulting from classic cartel behavior (firms meet secretly in hotel rooms to discuss prices and the consequences of cheating), or various cases in between. The harm from price coordination depends most directly on the extent and duration of supracompetitive pricing, not on the means of reaching or maintaining an understanding to charge the heightened price.

An economic approach to addressing coordinated oligopolistic price elevation through legal liability—like the economic approach to law enforcement more generally (*see, e.g.,* Polinsky & Shavell 2007)—seeks to determine liability and apply sanctions based primarily on the deterrence benefits that result as well as any chilling of desirable behavior that may ensue, while also considering the costs of operating the regime. Section 2.2 explores the problem of detection, the greatest challenge in the regulation of interdependent oligopoly pricing. Firms naturally seek to hide illegal aspects of their behavior, and reliable indicators are not always readily obtainable at low cost to enforcers. One set of detection techniques involves the use of market-based evidence to infer successful oligopolistic coordination, such as looking for sharp price drops associated with price wars. Others rely on internal evidence, which may pertain to firms' own estimates of costs and demand conditions as well as their understanding of their strategic situation vis-à-vis their competitors. Interfirm communications, when evidence thereof is available, also provide useful, perhaps decisive, information on whether successful coordination has occurred.

The degree to which conditions are conducive to coordinated oligopoly pricing is also quite relevant. Highly conducive conditions make inferences of successful interdependent pricing more credible—but hardly certain, among other reasons due to the possibility of effective deterrence—whereas unconducive conditions cast doubt on its plausibility. As will be explored in part 3, this

5 Exceptions are F.M. Scherer's (1977, 981–984) review of the first edition of Posner's book, which devotes a handful of pages to the subject at hand, and Richard Markovits, on which see note 9.

feature of sound detection strategy differs importantly from what many associate with a focus on the existence of particular interfirm communications because, under certain assumptions, more conducive conditions reduce the likelihood that such communications occurred even though they increase the magnitude of the net expected social harm from a failure to apply sanctions.

The analysis of sanctions in section 2.3 concentrates primarily on deterrence. A major challenge in setting monetary sanctions is determining the extent of price elevation, although this magnitude will often be indicated by much of the evidence on detection considered in section 2.2. In principle, the measurement problem is the same whether price elevation was accomplished through secret meetings, mere recognition of interdependence, or in any other manner. The threat of imprisonment as well as fines assessed against individual actors can be a useful supplement, particularly in light of agency problems within firms. Injunctions are also considered. Although much academic commentary fixates on injunctions, it is not evident that they are important in controlling coordinated oligopoly pricing.

Part 2 would seem to cover all pertinent aspects of the regulation of price fixing by considering the nature of the problem, how to detect its presence, and what remedies to apply. In the course of this investigation, the commonly advocated approach—which most commentators also contend describes current law in leading jurisdictions—of attacking only express and perhaps also tacit agreements, variously defined, barely surfaces. That is, it does not emerge from a systematic consideration of how best to address coordinated oligopolistic price elevation. The core of this article, part 3, systematically compares the communications-based prohibition and the direct approach presented in part 2 in order to ascertain whether anything important was overlooked and to identify explicitly the defects of the consensus method.

Initially, the conventional prohibition is defined, which turns out to be a difficult task.⁶ Most views can be captured by supposing that the price-fixing prohibition is limited to certain sorts of interfirm communication, whether designated by mode, content, or otherwise. This method on its face seems problematic because it focuses not on whether the means employed in fact caused harm in a given case but rather on whether one versus another means was employed. Moreover, such a distinction suggests that detection will often prove difficult, as indeed is the case.

The essential contrast with the direct approach is that the communications-based prohibition uses a large portion of the most relevant evidence indicative of undesirable behavior in an indirect fashion and also counts evidence

6 See Kaplow (2011c) for an exhaustive analysis.

concerning conduciveness of conditions backwards, in what is referred to here as a paradox of proof. Specifically, under certain assumptions that many endorse, evidence of an unusually high danger of successful coordinated oligopoly pricing exonerates firms instead of raising the likelihood that they will be subject to sanctions.⁷ As a consequence, the communications-based prohibition tends to assign liability in cases involving both less social harm and greater chilling costs compared to those targeted by the direct approach. The direct method, with an appropriately calibrated burden of proof, can be shown to dominate the communications-based prohibition in that, for a common rate of application, the direct approach targets cases with both relatively larger deterrence benefits and relatively smaller chilling costs.

These problems relating to detection and others arise because a communications-based prohibition requires that one not only determine the presence of interdependent oligopoly pricing—the focus of the direct approach elaborated in part 2—but also identify the means by which it was accomplished. This supplemental requirement is particularly difficult to meet because firms attempt to keep their methods secret, because the difference between permitted and prohibited means is formally rather than functionally determined, and because there is little empirical evidence that could guide the necessary inference process. As a result, the conventional approach is significantly more challenging to apply, which is ironic in light of its widely being favored on administrability grounds.

Part 4 considers a number of additional subjects: an alternative approach under which liability cannot be based on circumstantial evidence, contrary to the longstanding norm in competition law; implications of the contrasting approaches for other areas of competition law, such as the stringency of limits on horizontal mergers and the regulation of practices that might facilitate oligopolistic coordination; the manner in which rapid evolution in communications technology might influence the analysis, particularly concerning detection; whether dispositive motions can be helpful in quickly and cheaply disposing of weak cases; and how the various approaches square with legal doctrine.

Finally, the leading three arguments offered in favor of the traditional view are related explicitly to the foregoing analysis. One argument asserts a difficulty in attacking purely interdependent behavior because such would involve commanding firms to behave irrationally. This criticism is mistaken because it omits consideration of deterrence: applying heavy sanctions to certain choices will

7 Under other assumptions, which also find support, this paradox is reduced or eliminated, but in a manner that undermines most of the basis for paying attention to the conventional method.

change what firms find it rational to do. Another objection is that making price elevation by oligopolists illegal is inconsistent with the legality of price elevation by monopolists. This point ignores the purpose of separate, more stringent prohibitions on group behavior and, moreover, implies that classic cartels should be legal. Third, it is argued that remedies, particularly injunctive relief, directed at price elevation are problematic because they amount to price regulation. This claim is misconceived because, as mentioned, effective control is best accomplished through penalties that achieve deterrence rather than by relying on directive legal commands. However, this argument is suggestive of an important concern with competition policy regarding price fixing that is underdeveloped in the existing literature, namely, that the detection of violations can be quite difficult, raising the problem of false positives, the prospect of which chills desirable behavior. As emphasized throughout part 2, this concern should indeed be central in devising price-fixing rules, but the analysis here shows that it does not imply the desirability of the conventional approach over a direct method that takes explicit account of possible chilling costs.

The conclusion to the article emphasizes two respects in which the present analysis is incomplete. First, the focus throughout is not on the question that has preoccupied much prior discussion, “How should we define the term ‘agreement’?” but instead on the question, “What approach toward coordinated oligopoly pricing best promotes social welfare?” In answering the latter, it is natural to proceed by examining the nature of the problem and then determining how to identify its presence and to remedy it. Modern competition law emphasizes economic substance over form, has an open-ended, flexible formulation, and could in principle be amended. Also, under the laws of many jurisdictions, including the United States, a more substantive approach arguably conforms better to the statutory language, much of the relevant precedent, and aspects of existing practice than does the more formalistic approach that is widely endorsed—points developed in detail in a companion article (Kaplow 2011c). Even if prevailing doctrine imposes significant constraints, it is best to start by trying to determine what in principle is the most sensible way to address coordinated oligopolistic price elevation.

Second, optimal policy depends greatly on empirical evidence in realms where existing knowledge is incomplete. One set of issues concerns the extent of oligopolistic price elevation that would prevail under various regimes. Another involves the manner in which such coordinated pricing is achieved—for example, with resort to what sorts of communication—and, more broadly, how much of it can be detected, by which methods, and at what error cost. Without further knowledge, it is difficult to identify an optimal regime with confidence. However, the present framework not only guides that

decision in the interim but also sharpens the research agenda so that better strategies might be devised in the future.

2. DIRECT APPROACH⁸

A direct approach to the challenge of coordinated oligopolistic price elevation consists of identification of the social problem, detection of its presence, and application of appropriate sanctions. These three subjects are addressed in turn. The primary difficulty is in distinguishing successful coordinated price elevation from ordinary competitive behavior, so that actual violations can be sanctioned sufficiently often to achieve deterrence without unduly subjecting competitive behavior to the risk of sanctions that produces concomitant chilling effects.

2.1. Social Problem

Subsection 2.1.1 briefly reviews aspects of the theory of coordinated oligopoly behavior that are most relevant for the analysis of detection. Subsection 2.1.2 addresses the welfare implications of oligopoly pricing. Static effects are fairly familiar but dynamic effects are less well known and more subtle, and both are important in understanding the possible adverse consequences of false positives. Subsection 2.1.3 presents a framework for assessment.

2.1.1. Coordinated Oligopoly Behavior⁹

Economists study coordinated oligopoly behavior using the theory of repeated games.¹⁰ This theory aims to explain and predict oligopoly behavior that had previously been the subject of rough, intuitive, and not entirely satisfactory accounts (*see, e.g.,* Bagwell & Wolinsky 2002, 1872–1873;

8 For further exploration, see Kaplow (2011a).

9 As the section heading signifies, the analysis addresses coordinated oligopolistic price elevation. It has long been understood that, under certain conditions, even if the firms are unable to coordinate, price will exceed marginal cost (even though it will continue to fall short of the monopoly level). *See, e.g.,* Friedman (1986, 54–57); Kaplow & Shapiro (2007, 1083–1086); and Shapiro (1989, 333–356). For discussion of the implications of unilateral exercises of market power for policy toward coordinated oligopolistic price elevation, see Kaplow (2011a, sec. V.C). Unilateral price elevation is also explored (although with a different focus) in Markovits (1974, 1975).

10 The seminal article is Friedman (1971). On modern game theory in general, see Fudenberg and Tirole (1991) and Osborne and Rubinstein (1994); on the application to oligopoly theory, see Friedman (1986), Shapiro (1989), Tirole (1988, ch. 6), and Vives (1999, 301–323); and on the connection to competition law, see Kaplow and Shapiro (2007, 1103–1121) and Whinston (2006, ch. 2).

Friedman 1971, 11).¹¹ The relevant branch of the theory addresses what are termed noncooperative games, which are those in which binding agreements (such as might be enforced by a third party) are unavailable.¹² Note that this branch encompasses classic, explicit cartels (even written agreements) since competition law renders them legally void; parties must rely on themselves to enforce their arrangements. Unfortunately, the terminology can be confusing. Within noncooperative *games*, it is common to refer to both cooperative and noncooperative *outcomes*, a distinction that more closely tracks broader usage. Successful coordinated oligopoly pricing is a cooperative outcome; rivalrous competitive pricing is a noncooperative outcome (*see, e.g.*, Friedman 1986, 20; Fudenberg & Tirole 1991, xviii; and, for further discussion, Kaplow 2011c, sec. IV.A.1). Importantly, cooperation in this sense describes the character of the result, not the process by which the parties' minds (speaking metaphorically for firms) came to their mutual understanding. Thus, the theory focuses on what price levels can be sustained, using what forms of punishment and so forth, making no distinction based on what means of communication the parties may have employed. That is, the same analysis is applicable to classic cartels and to the most informal interdependence (*see, e.g.*, Kühn & Vives 1995, 43; Hall 2007, 1067).

The analysis of firms' interactions employs repeated games because these can capture the sort of strategic interaction that makes successful coordinated oligopoly pricing possible. It is familiar that, otherwise, price competition will yield the competitive outcome in a simple setting involving firms selling homogeneous products because firms that charge above marginal cost will be undercut by rivals.¹³ Unfortunately for society, firms sometimes are able to charge supracompetitive prices, despite the absence of legally binding agreements.

The basic idea is that starting, say, at the industry-profit-maximizing (monopoly) price, P_m , no firm will wish to cut its price to steal its rivals' business if it expects this act to induce its rivals to quickly cut their prices as well, perhaps matching the first firm's lower price and perhaps undercutting it. Whether, when the dust settles, the price is somewhat lower than P_m or all the way down to the competitive price, P_c , the prospective price cutter will be worse off. As long as it does not expect to profit sufficiently in the short run (before

11 Although the underlying story has long played a prominent role in competition law commentary and court opinions, the modern theory that makes it more precise has been virtually absent in these arenas. *See* Kaplow (2011c, 778–779 n.224).

12 Cooperative games are those in which externally enforceable binding agreements are possible, and the analysis focuses on what agreement parties would reach.

13 For more formal discussion that raises subtleties not central to the present task, *see, for example*, Shapiro (1989, 344–346) and Tirole (1988, 212–218).

rivals cut their prices as well) to make up for the sacrifice of profits in the long run, it will adhere to the price of P_m . Note that this logic applies regardless of whether each firm's expectation about others' reactions arises from their mutual appreciation of the situation or as a consequence of direct discussion of the matter.

Similar logic can explain how the price might rise to P_m in the first place, whether it starts at P_c or some intermediate level. A firm may be willing to brave a price increase if it expects (again, whether by conjecture or as a result of explicit discourse) that its rivals will reciprocate. If its rivals indeed cooperate by matching the price increase, the firms will all be better off thereafter, supposing that, by the logic of the preceding paragraph, the price increase can be sustained. As long as the first firm does not lose much profit in the interim (due to any delay in others' reactions), the long-run gain will make the venture worthwhile. Moreover, the firm will expect its rivals to follow quickly because they understand (again, either because of their grasp of the circumstances or through prior dialogue) that delay will be taken as defection, leading the initiator quickly to drop its price back to the preexisting level.

Modern game theoretic literature makes the foregoing intuition more rigorous and extends it along a number of dimensions.¹⁴ Perhaps the most important elaboration for present purposes concerns the problem that oligopolistic firms face in detecting defections in markets where each firm's prices are not readily observed—a challenge first explored in depth by Stigler (1964) (*see also* Spence 1978). In this setting, firms might be able to infer that someone has cheated from a fall in sales of their own products, and in simple models this information would be sufficient. But suppose as well that buyers' demand fluctuates in ways firms cannot observe, raising the possibility that a decline in a firm's sales might have been due to cheating or instead to less favorable market conditions. In the presence of such uncertainty, firms need to choose a strategy that trades off rapid, sufficiently harsh punishment of actual cheating—in order to deter it effectively—and avoidance of price wars when there was no actual cheating but just a period of unusually low demand. In models of this problem, coordinated oligopoly pricing may still be possible, but it is less effective; there will be occasional price wars even when no cheating has occurred. It has also been suggested that this more complicated scenario accords with what has actually occurred in some markets characterized by coordinated pricing.¹⁵

14 See generally sources cited *supra* note 10.

15 See, e.g., Green & Porter (1984); Porter (1983a, 1983b); Baker (1989); Bresnahan (1987); Levenstein (1997); and Shapiro (1989, 373–379); *see also* Perloff, Karp, & Golan (2007, 104) (finding the evidence more mixed) and Slade (1990) (examining alternative price-war models).

This extension and many others show how repeated games can be used to provide more complete depictions of oligopolistic markets, which aids in examining such questions as whether conditions are conducive to successful coordinated oligopoly pricing and whether it is occurring—both important inputs into inferences about the existence of successful oligopolistic coordination, the subject of section 2.2. Observe, as has been noted throughout, that there is no direct relationship between the analysis in this subsection and whether firms' success is achieved through pure interdependence or highly explicit communications. Both extreme, old-fashioned cartels (stripped only of the ability to legally enforce their agreements) and plain interdependence, along with everything in between, are analyzed as noncooperative games.¹⁶ Whether a pair of strategies constitutes an equilibrium for two gas stations engaged in price signaling with their price postings or having a discussion in a smoke-filled room leading them to charge the monopoly price depends on precisely the same calculation that compares the gain from defection with the lost future profits due to the other firm's response (*see* Whinston 2006, 21). Firms are assumed to act in their own best interests, to maximize profits, in either case. If the gains from cheating exceed the costs, it is supposed that a firm will cheat, and otherwise not.¹⁷

2.1.2. *Social Welfare Effects*

Just as with the descriptive theory of coordinated oligopoly behavior, implications for social welfare do not fundamentally depend on the means by which price elevation is accomplished, that is, whether it arises purely through recognized interdependence or results from secret meetings at which detailed plans are formulated. This rough equation is not particularly controversial; the point is often acknowledged by those who favor a narrow price-fixing prohibition (*see, e.g.,* Areeda & Hovenkamp 2003, 226). As will be seen, although some interesting differences do exist, they do not constitute substantial bases for limiting the social concern about price fixing to that achieved by particular means.

The core problem is that prices are higher than the competitive level. Such high pricing reduces both consumer welfare and overall efficiency (total welfare), the latter because choked-off sales are ones for which consumers' willingness to pay exceeds marginal cost, resulting in what is referred to as

16 For further elaboration, see Kaplow (2011c, sec. IV.B).

17 For qualifications and comments on their relevance, see Kaplow (2011c, sec. IV.B.3).

deadweight loss.¹⁸ It is important for subsequent analysis to note an obvious point: greater price elevations tend to be more harmful than small ones. This point is especially true if the objective is total welfare because deadweight loss rises disproportionately with price (the marginal loss equals price minus marginal cost, so this magnitude is greater when price is higher).

Static effects of oligopolistic price elevation are not necessarily limited to allocative inefficiency (deadweight loss) and transfers from consumers to producers. When oligopolistic firms raise price and accordingly reduce output, their output may not be allocated efficiently among them. Note in contrast that well-organized cartels can avoid this loss in production efficiency if they let their more efficient members produce more in exchange for transfer payments (Schmalensee 1987a). Accordingly, if oligopolistic price elevation is to occur to a given extent, classic cartels may entail lower efficiency costs than those that arise under pure interdependence or other less formal schemes.

Oligopolistic price elevation and deterrence thereof may also have important dynamic effects. As with static consequences, these effects do not as a first approximation depend on the means by which coordination is achieved. The dynamic effects considered here pertain to entry and inducing investment more broadly.

Regarding entry, the literature identifies two competing effects (*see, e.g.,* Mankiw & Whinston 1986). First, to the extent that prices are elevated above marginal cost, there tends to be too much entry due to a business-stealing effect: firms that enter obtain profits in part by diverting customers from incumbents, whose lost surplus is not taken into account by entrants.¹⁹ Therefore, resources wasted on excessive entry are an additional cost of price elevation. Importantly, this effect is the only pertinent one with homogeneous goods industries, and successful coordinated price elevation tends to be feasible primarily in industries with goods that are homogeneous or nearly so (a point consistent with the

18 For elaboration on the choice between consumer and total welfare as the appropriate criterion, see Kaplow (2011b) and Kaplow and Shapiro (2007, 1165–1169) (examining the choice with regard to the use of the efficiencies defense in horizontal merger assessment). Note that, even if consumers are to be favored on distributive grounds (perhaps because they are on average less well off than beneficiaries of firms' profits), it does not make sense to ignore producers' surplus entirely, and in any event it tends to be advantageous to achieve distributive objectives more directly, through redistributive taxation and transfers. *See, e.g.,* Kaplow & Shavell (1994).

19 There is a qualification in cases in which industry demand is insufficient to support even one firm or only one firm, although permitting effective monopoly pricing is not a good solution. Moreover, this problem does not directly relate to whether a price-fixing prohibition should be limited to that accomplished through certain means of communication. For elaboration, see Kaplow (2011c, sec. II.B.2.a).

facts of successful price-fixing prosecutions²⁰). Hence, dynamic effects on entry constitute a further social cost of coordinated oligopolistic price elevation in most cases. Also, like static costs, it is one that rises with the extent of the price elevation.

Second, to the extent that products are differentiated, there tends to be too little entry because firms offering new products do not capture all of the surplus generated by their contribution to product variety: firms consider only the revenue they obtain, ignoring inframarginal consumers' surplus. If this effect is sufficiently large, the additional entry induced by price elevation is socially beneficial. However, a substantial benefit from enhanced variety arises only when differentiation is great, and in such industries coordinated price elevation is unlikely to occur.

Entry is just one form that investment may take. Viewed more broadly, elevated prices tend to reward whatever activities lead firms to be in a position to charge prices in excess of marginal cost or to capture a larger share of the market when doing so. That the allocative inefficiency of supracompetitive pricing may often be outweighed by the dynamic gains from investment induced by the prospect thereof constitutes the classic justification for granting intellectual property rights. More broadly, competition regimes, such as that in the United States, tend to tolerate monopoly pricing as long as the monopoly position was obtained and is maintained through efficient behavior—producing products that consumers value and selling them at sufficiently attractive prices—rather than through exclusionary practices.²¹ This permissive approach is not without costs; ex ante incentives can be excessive, some ex ante investment is wasteful (rent seeking), and ex post payoffs, with accompanying allocative inefficiency, may be greater than is necessary to induce investment. Hence, intellectual property rights tend to be limited, and direct regulation of monopoly is sometimes imposed. However, the judgment behind general competition regimes is that, in other settings, monopolies should be permitted to price as they wish, although the degree to which this view is accepted varies across jurisdictions.

Price elevation achieved by interdependent behavior among oligopolists is qualitatively different in this regard. Such price elevation does not reward firms

20 Examination of the lists of industries in Connor (2007, 136–153), Harrington (2006, 98–102), and Hay and Kelley (1974, 29–38; also 24–25, 27) strongly suggests this to be true, and it is certainly true of the large international cartels examined in Connor (2008).

21 See, e.g., *United States v. Aluminum Co. of America*, 148 F.2d 416, 430 (2d Cir. 1945) (“The successful competitor, having been urged to compete, must not be turned upon when he wins.”); *United States v. Grinnell Corp.*, 384 U.S. 563, 570–571 (1966) (referring to “the willful acquisition or maintenance of that [monopoly] power as distinguished from growth or development as a consequence of a superior product, business acumen, or historic accident”).

to the extent that they outperform their competitors but instead bestows profits whose magnitude depends on firms' success in refraining from competition.²² To be sure, the prospect of such rewards will induce a variety of ex ante investment behavior. In addition to entry, already discussed, firms might expand their production capacities if they expect their capacities to have a positive effect on their share of oligopoly rents, and this expansion would be wasteful if the additional capacity would remain idle. Similarly, some activities, like advertising, may be zero-sum (to an extent). However, other types of investment are socially valuable, such as product improvements and cost reductions. Yet, even regarding the more useful activities, the reward from oligopoly pricing will substantially reflect the firms' ability to abstain from competition rather than the merits of their prior investments. Thus, although permitting oligopolistic price elevation may produce some dynamic gains, there is insufficient nexus to justify price fixing (*see* Schmalensee 1989, 989).

Furthermore, oligopolistic pricing reduces potential dynamic benefits. When some oligopolists are more efficient or offer products superior to those of others, successfully coordinated pricing tends to dampen the tendency of better firms to serve an increasing share of consumers. In addition, firms may have less incentive to become more efficient and innovative in the first place because they may not greatly benefit from such activity unless they will be willing to defect from the interdependent arrangement, which they will be reluctant to do if the oligopoly profit margin is substantial. Also, economies of scale are not fully realized by successful oligopolists. Competition laws may allow competing firms to merge or enter into joint ventures precisely because of the potential to realize efficiencies; however, such actions are subject to review in order to verify that anticompetitive effects are negligible or perhaps are justified in light the efficiency gains. Coordinated oligopoly pricing incurs the costs without producing these benefits. Finally, note once again that little of the foregoing analysis suggests any direct distinction between unaided, albeit successful oligopolistic interdependence and price elevation achieved through explicit communication.

Some commentators have argued that plain interdependent oligopoly pricing should be permitted because price elevation by monopolists

22 In this regard, unilateral market power by oligopolists must be distinguished. *See supra* note 9. Demsetz (1973, 3) offers a classic statement of the ex ante investment benefits of the prospect of market power and their inapplicability to coordinated oligopolistic pricing:

To destroy such power when it arises may very well remove the incentive for progress. This is to be contrasted with a situation in which a high rate of return is obtained through a successful *collusion* to restrict output; here there is less danger to progress if the collusive agreement is penalized.

is legal.²³ This claim is surprising. Competition law explicitly distinguishes and subjects to tough sanctions the efforts by groups of firms to eliminate competition among themselves, in contrast to the price-elevating behavior of monopolists. Moreover, this differential approach is well founded, as just explained. A final reason the argument is puzzling is that the same logic would allow classic cartels—indeed, even legally enforced cartel prices, for a monopolist can legally force its employees to charge the price it commands—yet those advancing the argument roundly condemn express price fixing.

2.1.3. Framework for Decision-making

If detection of successful oligopolistic interdependence was costless and perfect—that is, if all such behavior was detected and no innocent behavior was misclassified—then the only question would be how to set sanctions optimally, the subject of section 2.3. A substantial challenge remains because detection of successful coordination is inevitably imperfect. Accordingly, a central question—probably the most difficult and important one regarding price fixing—is setting optimal proof burdens, that is, deciding how much of what sorts of evidence in various contexts should be deemed sufficient for a finding of liability.²⁴

First, consider situations in which firms have in fact engaged in coordinated oligopolistic price elevation. In such cases, the primary benefit of assessing liability is deterrence; thus, the social cost of false negatives—failures to identify price elevation that has in fact occurred—is the loss of deterrence. If it is possible to raise sanctions to make up for the detection deficit, there would be no difficulty. There are, however, often limits to how high sanctions can be (firms will be judgment proof beyond a certain point, for example), and there are also costs in trying to identify and prosecute acts of price elevation. Therefore, it seems likely that insisting on greater certainty of proof will involve some loss in deterrence due to the larger portion of false negatives that result. Of additional concern, the failure to detect oligopolistic price elevation will not be random. If some methods of proof are allowed but others are restricted (they may be disallowed or subject to high proof standards), it may well be that certain groups of firms will be deterred and others not. Further raising the magnitude of sanctions may primarily relate to the former, where there may be little or no

23 See, e.g., Areeda & Hovenkamp (2003, 232) (arguing against prohibition of pure interdependence because such would be inconsistent with rules on monopoly); *id.* at 272 (same); and Turner (1962, 668) (“It would make no sense to deprive lawful oligopolists—those who have achieved their position by accidental events or estimable endeavor—of the natural consequence of their position if the lawful monopolist is left with his.”). *But see* Lopatka (1996, 854–855) (criticizing this view).

24 On the general question of how to set proof burdens optimally, see Kaplow (2012a, b).

deterrence deficit, while having little effect on the latter. This point is particularly relevant in assessing communications-based prohibitions that *de facto* or *de jure* exonerate oligopolistic price elevation in certain settings. Obviously, raising penalties will not deter groups of firms that are effectively immune from them.

Second, consider firms that have not engaged in oligopolistic price elevation. The prospect of sanctions may tend to chill (deter) beneficial activity in settings in which firms anticipate that their actions generate a substantial risk of false positives.²⁵ It is familiar that, if certain medical procedures involve a high risk of malpractice liability even when doctors behave properly, doctors may refrain from such procedures even when they are beneficial. Commentators, courts, and enforcement agencies that have been reluctant to take too aggressive an approach toward price fixing seem implicitly to have this sort of concern in mind. The point is especially significant if evidence that firms engaged in similar or identical pricing behavior is taken to be proof of price fixing, or close to it, for such behavior is the norm even among perfect competitors. Logically, such evidence does not even begin to make the case for coordinated oligopolistic price elevation. What is required is to examine a range of evidence that is indeed probative, taking advantage of multiple methods and using them in complementary ways.

The cost-benefit calculus in setting proof burdens involves a tradeoff of deterrence benefits and chilling costs. Demanding more compelling proof as a predicate for assigning liability will tend to reduce both deterrence benefits and the costs of chilling desirable behavior. Likewise, greater openness to less definitive proof will enhance deterrence and amplify chilling effects. However, the problem is more complicated because different channels of proof vary substantially in their contribution to deterrence and their risk of chilling effects. Ultimately, approaches toward different sorts of evidence need to be functional, considering their benefits and costs directly, rather than employing formalistic criteria that arbitrarily limit modes of proof or alter the target of inference in a way that deviates from the social objective.

A direct approach assesses the importance of errors, both false positives and false negatives, in terms of their effects on social welfare. More conventionally, errors are often understood by reference to a formal legal criterion, even when that criterion itself is chosen because it is a proxy indicator of which behavior should be sanctioned.²⁶ Most relevant for present purposes, the application of

25 False positives may also reduce the deterrence of coordinated oligopolistic price elevation because the incremental difference in expected sanctions from engaging in coordinated pricing versus abstention therefrom is reduced.

26 Another important difference is that errors are sometimes conventionally viewed as if they are intrinsically bad, and the manner of specifying their weight is mysterious.

punishment in a case in which no prohibited communications are demonstrated may be regarded as a false positive. However, if certain classes of such supposedly erroneous findings entail the assignment of liability in settings in which coordinated oligopolistic price elevation in fact occurred—but without the requisite communications, or at least without proper proof thereof—then these so-called false positives would be desirable, not detrimental, in terms of their effect on social welfare. Clearly, the prospect of such false positives that are defined by reference to the formal legal rule will deter undesirable behavior, not chill desirable behavior.

2.2. Detection

It is important to consider how detection would be undertaken following a direct approach, both for its own sake and because this assessment will sharpen the contrast with a communications-based prohibition in part 3. A direct inquiry asks whether coordinated oligopolistic price elevation²⁷ has occurred and not whether, if it has, it arose as a consequence of one or another means of interfirm communications²⁸—although, as will be discussed, such communications can be significant evidence bearing on whether successful coordination has occurred. As noted, the social consequences of interdependent pricing do not depend on this matter. Whether some methods of proof are to be privileged above others should reflect analysis of the pertinent inferences and application of the appropriate decision-making framework. Accordingly, any special or exclusive relevance of interfirm communications, or any other indicator of coordinated pricing for that matter, is not determined a priori but instead needs to emerge from the functional analysis.

27 The emphasis is on coordinated elevations rather than on those that arise from the unilateral exercise of market power, on which see note 9 and, for further exploration of its relevance to the present problem, Kaplow (2011a, sec. V.C).

28 See, e.g., Jacquemin & Slade (1989, 452) (in discussing certain economic methods of identifying collusion, they state: “[I]t is impossible to distinguish pure tacit collusion from . . . explicit cartel agreements. What matters for the empirical estimates is the outcome and not the cause of non-competitive pricing.”) and Porter & Zona (2008, 1071) (“As a matter of economics, it is difficult and perhaps impossible to distinguish between [interdependent oligopoly, conscious parallelism, tacit collusion, and explicit collusion] on the basis of outcomes alone.”). It is sometimes suggested that one occasionally can tell whether elaborate, explicit communications have occurred, particularly when behavior is unusually sharp and precise (e.g., secret bids that are identical down to many digits). Even this point is overstated, for often the opposite inference might be made instead because, the more explicit were the communications, the more readily firms could have orchestrated their behavior so as to avoid leaving clear, visible tracks. See *infra* note 65.

2.2.1. *Market-Based Evidence*

2.2.1.1. *Means of Inferring Successful Oligopolistic Coordination.*—There are myriad means of inferring the existence of successful oligopolistic coordination,²⁹ many of which are not independent of each other: some may be simultaneously present or mutually reinforcing. A converse is likewise important: regarding some factors that imply coordination, their absence may often negate its existence, where absence should be understood as not merely constituting ambiguity and difficulty of proof but rather demonstration of nonexistence.

Despite the overlap and interrelationship, it is useful to group indicators into two clusters: pricing patterns and price elevation. Certain pricing patterns may indicate successful oligopolistic coordination or a breakdown that implies its prior existence. Given the long history of confusion on this subject, it is best to begin by emphasizing that the presence of parallel pricing and other shared behavior is not usually a symptom for the simple reason that ordinary competitive interaction also has this character. Indeed, when competition is vibrant, most pricing and other behavior are parallel. When firms' costs increase, their prices rise and quantities fall. When demand increases, firms' prices and quantities both rise. When technology changes, consumers' locations or tastes shift, regulations are modified, and so forth, competitors react similarly, even identically. The goal is to distinguish successful oligopolistic interdependence from competitive, independent, rivalrous behavior. As a logical matter, traits, such as parallel pricing, that are shared by both categories are not useful in drawing the distinction. Instead, analysis should focus on behavior that is consistent with oligopolistic interdependence and inconsistent with competition, which favors liability, and behavior consistent with competition but inconsistent with interdependence, which disfavors liability.

In considering pricing patterns that may support inferences of successful oligopolistic coordination, it is useful to analyze three phases: raising prices from a competitive to a supracompetitive level (or further escalating prices), maintaining elevated prices, and price drops, notably, as a consequence of price wars. Initiation or enhancement of oligopolistic price elevation may be marked by a sharp price increase. However, such price increases can also happen in competitive markets, most obviously when there is a cost shock, such as a sudden increase in the price of oranges for grocery retailers or in the price of oil for sellers of gasoline. Accordingly, it is also necessary to check for concurrent cost increases or other changes, such as sudden shifts in demand, that may explain the price increase. Additionally, signaling and jousting about price—such as

29 Prior treatments include Baker & Bresnahan (1992, 2008), Bresnahan (1989, 1997), Harrington (2008), Kaplow & Shapiro (2007, 1087–1095), Perloff, Karp, & Golan (2007), Porter (2005), and Posner (2001, 79–93).

through sequential price changes, where firms await others' reactions and then proceed—may sometimes provide a basis for inferring oligopolistic elevation. A difficulty is that sophisticated firms, aware of what inferences may be drawn from their price moves, may instead adjust prices strategically in order to disguise their coordinated behavior. Observe that more complex behaviors of this sort might require more elaborate (and thus more explicit) communications that may leave traces. Note that, if such camouflaging strategies are net helpful to firms, an implication is that interdependent behavior involving little or no direct interfirm communication may be easier to distinguish from competitive behavior than is interdependent behavior implemented after more explicit interchanges.

Another familiar point is that coordinated prices, and perhaps also market shares, will tend to be stickier over time than are those of competitors. The main reason is that, because coordination is difficult, frequent fine-tuning may be avoided. In contrast, competitors do tend to respond quickly, and if changes are frequent, often. There are other explanations for sticky prices, such as menu costs (referring to the cost of changing prices per se), but in some settings such costs are negligible or insufficient to explain the extent of stickiness.

Sudden, sharp price reductions are as suspicious as sudden, sharp price increases, again, in the absence of corresponding changes in cost or demand. Oligopolists do not ordinarily wish to drop the price, but such is sometimes unavoidable. Price wars arise to punish cheaters or when firms experience a loss of customers and thus must act on the assumption that cheating occurred even though a decline in demand that is not yet evident might have been the cause (as explained in subsection 2.1.1). If prices fall precipitously, without any exogenous change of corresponding magnitude, the preexisting price probably involved coordinated elevation.³⁰

Another strategy to detect elevation is to examine prices themselves. A logically straightforward way to determine whether price exceeds a competitive level, marginal cost, is to compare price and marginal cost directly. While price is often easy to determine, marginal cost may be quite difficult to measure in many settings. This challenge is familiar in competition law enforcement, often being confronted when assessing market power³¹ and also at issue in predatory pricing disputes³² and in some other contexts. Underestimating marginal cost

30 The use of price cuts to enforce coordinated oligopolistic price elevation has motivated proposals to further take advantage of this feature. See Bishop (1983) and Sagi (2008).

31 If measuring existing (that is, exercised) market power, the problem is tantamount to measuring price-cost margins. See, e.g., Kaplow & Shapiro (2007, 1079–1080).

32 It is common to inquire whether price is below cost (although the appropriate notion of cost is not always defined). See, e.g., *Brooke Group Ltd. v. Brown & Williamson Tobacco Corp.*, 509 U.S. 209

and thus producing false positives can occur as a consequence of excluding costs as fixed that are really variable (perhaps equipment could be rented or sold), ignoring common costs (which may actually need to be raised to support additional output of the product in question), or failing to recognize that marginal costs may rise steeply when output is close to capacity. In response, enforcers could consciously err in the opposite direction (treating more costs as variable costs of the product in question, for example) and require demonstration of a more significant elevation of price above measured marginal cost.

Another approach relies on comparing prices for the same product sold to different purchasers. Prices might be compared across markets; notably, geographical price differences may indicate supracompetitive pricing in regions with higher prices. Similarly, different prices charged to different customers in a single market—that is, the presence of price discrimination—may indicate the existence of market power. Both methods have their limitations, such as the need to attend to possible cost differences.

A different way to identify successful oligopolistic coordination is to examine whether ordinary pricing behavior (that is, aside from episodes of commencing or terminating coordination) responds to changes in demand and cost in the same manner as would pricing by competitors or somewhat (or entirely) like pricing by a monopolist (*see, e.g.*, Bresnahan 1989, 1012; 1997, 71). A substantial body of econometric research since the 1980s is designed to measure the exercise of market power using such a strategy (*see sources cited supra* note 29). This work represents a significant advance in market power measurement, but it is hardly a panacea; the techniques often require strong assumptions about the structure of demand and cost that it may not be possible to test directly (*see, e.g.*, Perloff, Karp, & Golan 2007, 42, 70–71, 91).

The foregoing list is hardly exhaustive. It is clear that successful oligopolistic coordination resulting in price elevation has a number of features that distinguish it from ordinary competition and that there exist techniques for identifying these differences. On the other hand, these differences will not always be apparent, and sometimes alternative explanations will be available that themselves may be difficult to assess.

A different sort of evidence concerns the use of facilitating practices: acts that make it easier to engage in oligopolistic coordination (*see, e.g.*, Buccirossi 2008; Posner 2001, 86–87, 88–89, 91–93). Commonly discussed examples include advance price announcements, discussions at trade association meetings, product standardization and suppression of quality competition, cross-ownership,

(1993). Note that if true marginal cost is above measured cost, which as the text to follow indicates is the more likely error, predation tests are too lax but determinations of the existence of coordinated oligopolistic price elevation are too strict.

most-favored customer clauses, and price-matching guarantees. Indeed, explicit interfirm communications, taken by many to be determinative of what constitutes a price-fixing agreement, can also be viewed as facilitating practices (see Kaplow 2011c, sec. II.B.2).

Facilitating practices may be relevant in two ways.³³ First, their use provides a basis for inferring the existence of oligopolistic coordination. This inference is sensible when there exists no other plausible explanation for the practice. Second, facilitating practices may themselves be made a basis for liability. Again, one must determine whether there are other explanations for use of the practices, and, if there are, any benefits must be weighed against the competitive risks.³⁴

2.2.1.2. Conduciveness of Conditions.—An assessment of the degree to which industry conditions in a particular setting are conducive to successful oligopolistic coordination can sharpen the accuracy of inferences on the ultimate question of whether such coordination is occurring. This idea is especially important for present purposes because the relevance of conduciveness under a communications-based prohibition may be radically different—arguably the opposite in many settings, as elaborated in sections 3.4 and 3.5. Indeed, perhaps the greatest potential distinction between the two approaches toward price fixing concerns this divergence in the use of evidence on conduciveness.

To begin, conduciveness may be seen as a necessary condition for successful coordination. If coordinated oligopoly pricing is impossible under the circumstances, we can confidently infer that it is not taking place. Similarly, if conditions are unconducive, success seems unlikely.³⁵ A key qualification is that the investigator or adjudicator may be mistaken about how unconducive the conditions actually are. Firms ordinarily have a better grasp of industry conditions and of their own ability to succeed in spite of them than will an outsider. Thus, if one sees clear attempts to coordinate or strong evidence that oligopolistic coordination is successful, the better inference is that the mistake lies not with the firms' analysis of conditions but rather with the observer's.

33 The nature of facilitating practices is also relevant to whether conditions are conducive to collusion, the subject of the next subsection.

34 See also *infra* section 4.2 (discussing how a more limited prohibition on coordinated oligopoly pricing may require a stricter and thus more costly regulation of facilitating practices).

35 One reservation to this argument concerns ill-conceived attempts. It is often helpful to punish attempts even if they fail or, importantly, when it is difficult to know whether they have succeeded. Nevertheless, the deterrence benefit of such punishment is likely to be limited when considering realms in which success is difficult or impossible in any event, and we should be more skeptical that an attempt occurred in unconducive circumstances. See Kaplow (2011a, sec. V.B).

It should be clear as well that inferences about conditions should influence inferences about successful coordination and vice versa. If we are highly confident about one set of inferences, then we should adjust, perhaps significantly, our inferences about the other. As mentioned, if conditions seem quite unconvulsive, we should be more skeptical of evidence that may otherwise lead us to believe that oligopolistic price elevation took place. However, if evidence on the latter is quite strong, then we should doubt evidence that conditions are highly unconvulsive. Often, there will be nontrivial uncertainty surrounding both matters. In addition, even when there is reasonable confidence regarding convulsiveness, it may be that it is at an intermediate level, readily admitting the possibility that coordination would succeed and that it would fail. In such cases, we would need to rely primarily on the strength of the evidence bearing on success to determine whether oligopolistic coordination took place.

Although convulsiveness at some level is a necessary condition to success, it is not a sufficient condition, and this is so even when conditions are extremely convulsive. The primary reason is that deterrence may be effective. Even the easiest task will not be undertaken if the expected penalty for doing so is sufficiently great. In a well-functioning regime, coordinated oligopolistic price elevation may be deterred in most instances. Put another way, unless one believes that the legal system is a substantial failure, one cannot assume that highly convulsive conditions in and of themselves imply a significant likelihood of coordinated price elevation.

Convulsiveness is also important with regard to the magnitude of deterrence benefits and chilling costs. When conditions are quite unconvulsive, any successful oligopolistic coordination is likely to result in price elevations that are small and short-lived. In addition, chilling costs are likely to be greater in such situations because these industries are reasonably likely to be competitive. There are a large number of markets where conditions are quite unconvulsive to successful coordinated price elevation. If such cases are not screened out, they might give rise to many false positives. On the other hand, when conditions are highly convulsive, undeterred firms may succeed in significant, long-lasting price elevations, whereas chilling costs may be less worrisome since highly competitive behavior is less likely. Accordingly, insisting on moderately or perhaps highly convulsive conditions seems sensible.³⁶

The present discussion will not elaborate the conditions themselves because the subject has received extensive attention in the literature (*see, e.g.*, Connor

36 As will be discussed in sections 3.4 and 3.5 on the paradox of proof, some views of more traditional approaches to liability, which emphasize explicit interfirm communications, may have the opposite implication regarding convulsiveness, which for the reasons just given in the text serves to reduce rather than maximize social welfare.

2008, 32–42; Hay & Kelley 1974, 14–17; Kaplow & Shapiro 2007, 1108–1121; Posner 2001, 69–79; Vives 1999, 306–310) and is not particularly controversial. Moreover, the particulars are not very germane to the comparison of the two approaches (although they are pertinent to how both are implemented since conduciveness is highly relevant under both). Some of the familiar factors are numbers of firms and market concentration, product heterogeneity, the extent of asymmetry across firms, the transparency of prices, the structure of the buyer side of the market, demand uncertainty, and the coordinating group's market power.

2.2.2. *Internal Evidence and Communications*

For virtually any factor examined in subsection 2.2.1, there is corresponding internal evidence. In addition to hard evidence—for example, invoices or other records indicating prices charged and quantities sold—there is a variety of other information. Some of it may directly convey firms' thinking (strategy or decision memos, notes of meetings, internal policy pronouncements), and much more will convey aspects indirectly (cost and marketing data being suggestive of firms' beliefs about marginal cost and demand). After all, a substantial portion of higher-level managers' time, as well as efforts throughout particular departments, is devoted to information gathering, analysis, and decision-making. These activities generate all manner of notes, reporting up and down chains of command, memos, communications, meetings, and so forth. They relate to what firms know or believe as well as to the reasoning behind their decisions.

For example, when a firm suddenly increases its price, it will have reasons, usually supported by its own evidence. If there is a corresponding increase in cost, this will be reflected in the firm's internal cost data and projections as well as in its decision-making process. When price suddenly drops, similar internal information and activity will also be involved. Whether the firm is a secret price cutter, is responding to perceived cheating by others, or is reflecting changes in cost or demand, the information and rationale will be reflected internally. More broadly, if firms are interacting interdependently, we would expect internal discussions and other traces to differ from the situation in which firms behave as competitors who take rivals' behavior as given.³⁷ Note that omissions can also be revealing. For example, a lack of new data or of discussion about

37 For example, in some court cases (see Kaplow 2011c, 747–748 n.158), there is evidence of firms' "understanding" with others, viewing competitors as "friends," and so forth, all of which suggest interdependence. Interestingly, courts referring to such indicators tend to view them as evidence of express communications in contrast to pure interdependence, a logical error since pure interdependence involves a meeting of the minds and resulting mutual understanding that is equivalent to that existing in a classic cartel; the difference is in how that understanding comes about. See *infra* section 3.3.

changes in demand makes a sharp demand shift an unlikely explanation for a price move. In contrast, if a new product proposal is rejected because studies detail how the firm lacks production capability or how consumers are uninterested, suppressing heterogeneity is an unlikely explanation.

There are important limitations in relying on internal evidence from firms. First, since a large firm is a group of many individuals with complex, overlapping, and sometimes conflicting duties, it can be difficult to determine what a firm knows or what reasoning explains its actions. Many decisions are based on soft information or are made despite seemingly contradictory information, perhaps because the information is seen to be unreliable, because there are overriding considerations, or because of incompetence. There may exist internal disagreement or misunderstanding. Firm politics may also play a role, reflecting that employees are not perfect agents of the owners. A second concern is that lawyers or others will anticipate the liability implications of firms' knowledge, decision-making, and actions and therefore will sanitize or distort the various clues to be found within the firm.

Nevertheless, it seems plausible that some substantial and fairly reliable conclusions will often be possible to reach. Firms cannot operate while lacking even an approximate sense of their costs and market demand, without any rationale for their actions. If successful oligopolistic coordination significantly elevates prices for extended periods of time, there may well be substantial internal indicators. Similarly, if no such behavior has occurred, there may be evidence inconsistent with hypothesized coordinated price elevation. Relatedly, it is difficult for a team of lawyers or others to reach broadly and deeply into a large corporation, controlling how myriad individuals speak, write, email, and otherwise behave. Furthermore, any such process that is not limited to a handful of key agents may itself leave incriminating tracks. Also, even if no individuals made explicit reference to forbidden activity, most of the information considered here would remain, notably that concerning firms' costs, demand, strategic decision-making, and so forth. Firms' task at hiding their tracks is even more daunting because we are examining the possibility of coordinated behavior. Consider how difficult it would be for multiple firms to fabricate consistent stories, which could survive scrutiny, about a cost shock that had not in fact occurred. In contrast, if there really was a cost shock, and accordingly the firms' simultaneous price increase was an ordinary response of competitors, substantial corroborating material would likely have been generated in the ordinary course of business by each of the firms, and it would be consistent with external indicators.

Interfirm communications constitute an additional source of evidence when making inferences about whether successful oligopolistic coordination has taken place—that is, (even) under a direct approach that does not make liability

contingent on particular prohibited communications. Indeed, when documented and specific, they are often excellent evidence. Interfirm communications contribute to our understanding of what firms have done and why. Note that public interfirm communications, ranging from advance price announcements to commentary on industry conditions, are not hidden and may be as revealing as secret discussions.³⁸ Even when they are directed toward the public, sequential advance announcements arriving at a consensus price may indicate what firms are doing. In this respect, they may supplement information from price moves themselves; for example, if an industry-wide price increase is due to a common cost shock and firms are behaving competitively, they will have less of a need to feel out their rivals before raising their prices.

Second, interfirm communications can serve as an important facilitating device and thus support an inference of coordinated pricing. These might include public statements—advance price announcements, predictions about industry demand or costs, and open discussions of various matters at trade association meetings—as well as secret meetings. These communications may help firms reach a consensus and be especially useful when more complex understandings are required, perhaps because of product heterogeneity. In addition, direct interchange may help identify and coordinate the punishment of cheaters. Moreover, some forms of interfirm communications are an attractive enforcement target because there is little social cost if firms are deterred from engaging in them. Secret meetings in hotel rooms come to mind. For other communications, such as public price announcements, trade association activities, and participation in standard setting or joint ventures, there are varying degrees of possible benefits that need to be considered.³⁹

Consequently, interfirm communications of all sorts can be important in making reliable inferences. It is also clear that their reliability may be enhanced, perhaps significantly, by combining them with the other evidence considered throughout this part. Accordingly, evidence of interfirm communications should be viewed as a weapon in the detection arsenal but not as the be all and end all.

A related question is how probative is the lack of explicit, detailed interfirm communications. First, because firms attempt to keep these secret and may succeed in doing so, a failure to find such communications is hardly conclusive that they did not occur. Second, oligopolistic coordination is sometimes

38 Note that, if public interchanges were per se legal, firms could simply move their meetings from hotel rooms to joint press conferences.

39 In this regard, as section 4.2 explains, a legal regime that is more aggressive toward successful oligopolistic coordination may be able to be more relaxed about activities that have social benefits but also might facilitate price elevation (which can be downplayed if it is independently deterred).

possible without elaborate communications. Regarding both points, there is an important interaction with the conduciveness of conditions. When conditions are most conducive—particularly when the number of firms is small—explicit interfirm communications are both less essential and more difficult to detect.⁴⁰ This suggests that, when the danger of oligopolistic price elevation is greatest, we should be less bothered by the lack of such evidence. Similarly, when firms' coordination problem is highly complex and likely to require extensive, explicit communications among large numbers of individuals, the absence of direct (or strong indirect) evidence that any such communications took place would make an inference of successful oligopolistic coordination notably weaker.⁴¹

An additional point, which will be explored further in sections 3.4 and 3.5, is that it generally makes little sense in cases where no secret interfirm communications are directly detected to attempt to infer through the use of circumstantial evidence whether they have taken place. Since the reason we care about these communications is that they further illuminate other evidence, nothing is added by this process. Stated abstractly, if evidence E_1 , E_2 , and E_3 give rise to a probabilistic inference that internal communication C occurred, any implication from the set E_1 , E_2 , E_3 , and C about successful oligopolistic coordination can be no stronger than what one could have inferred directly from E_1 , E_2 , and E_3 alone.⁴² As we will see, the process of attempting to infer specific sorts of communication from circumstantial evidence is complicated and uncertain, and it generally entails having already reached a judgment on whether coordinated oligopolistic price elevation has occurred.⁴³ Thus, the supplemental inference about communication is of little use, very costly, and also highly distracting.

2.3. Sanctions

Fines are widely used to penalize price fixing, supplemented by damages in the United States and to a lesser extent in other jurisdictions. Moreover, basic law enforcement theory (*see, e.g.*, Polinsky & Shavell 2007) suggests that their use is often desirable, for monetary sanctions tend not to be costly in themselves (in contrast to imprisonment and injunctions), they deter behavior (which

40 See the elaboration and qualifications in subsection 3.4.5.

41 As will be seen in subsection 3.4.2, the logic in this paragraph differs greatly from that implied under the commonly advanced view.

42 Even when direct evidence of C would strengthen the ultimate inference, if the only evidence of C is due to the inference from E_1 , E_2 , and E_3 , nothing is added.

43 The ultimate irony arises when defendants grant this point—that is, that they are elevating price—and use the ease of coordination to negate the existence of particular sorts of communication. Such is the paradox of proof, elaborated in sections 3.4 and 3.5.

injunctions alone do not), and they may be calibrated to the extent of harm caused and the likelihood of detection. Regarding the latter, the general prescription is that, ideally, expected sanctions should be set equal to expected external harm, so fines or damages should equal actual harm multiplied by the inverse of the likelihood of successful detection. Because oligopolistic firms attempt to keep their price coordination secret, this latter point is important in the present setting, and this is the rationale for employing treble damages in private lawsuits in the United States—which, note, are in addition to any fines and prison terms imposed by government enforcers. Additionally, because chilling costs are of greatest concern when industries are actually competitive, which will tend to be situations in which apparent (although mistakenly identified) price elevations are small, correspondingly lower sanctions in cases with low measured price increments provide a useful if incomplete safety valve. Therefore, it is important to attempt to measure harm reasonably well when determining the magnitude of fines or other sanctions. Fortunately, much (but not nearly all) evidence used to detect successful oligopolistic price elevation also provides a basis for estimating its magnitude (*see, e.g.*, Baker & Bresnahan 1992, 15). And when it is necessary to inquire further into the magnitude of price elevation for purposes of setting the level of sanctions, it makes sense to use this information to refine inferences regarding liability.

Although the core deterrence logic is simple and familiar, it is worth emphasizing because it seems that many competition law commentators fail to appreciate it, as discussed below, despite Posner's (1969, 1588–1593) clear presentation of the basic idea in his seminal article decades ago. Whatever methods firms might use to achieve coordinated oligopolistic price elevation, firms' decision-making calculus changes if they expect to be subject to sanctions. If they anticipate a large gain from successful oligopolistic coordination and no penalties, they will expend great effort to accomplish it. However, if there are sanctions and their expected level exceeds expected profits from coordination, then they will voluntarily forgo such efforts and aim to ensure that their employees do not engage in this activity. The logic is no different for, say, the application of a corrective tax to firms that pollute. In imposing the tax, the government simply charges firms for their pollution; how pollution is reduced is up to the firms. With price fixing, the optimal degree is ordinarily zero, and if expected sanctions are equal to harm to others and thus exceed any profits, that is the level firms will choose.

The use of imprisonment is socially costly: resources are consumed in running the system and the loss to imprisoned individuals is not matched by any direct social offset (unlike fines, which are transfers). Accordingly, imprisonment is not optimally employed unless monetary sanctions are insufficient. Imprisonment, which is used in price-fixing enforcement, especially in the

United States, can nevertheless be important. Detection is difficult, which may require for adequate deterrence a level of fines that exceeds firms' assets, and firms cannot always readily control employees, some of whom may have incentives to violate the law (for example, to receive a bonus or obtain a promotion due to higher current profits) even when violations are against the interest of the firm as a whole.

Injunctions are notably different from fines and imprisonment. Most importantly, they do not by themselves achieve deterrence. Firms would be happy to elevate prices for as long as they could get away with it—which may be a long time given detection difficulties and also the time delay involved in adjudication—if the only cost upon apprehension is the need to abstain in the future. (If the only legal consequence of theft, murder, or tax evasion was the possibility of being ordered not to repeat the violation in the future, crime may become rampant.) Furthermore, if fines and damages (and imprisonment) are to be employed, why is there a further need for an injunction? After all, when the other sanctions succeed at deterrence, whether to impose injunctions becomes moot. Also, injunctions are parasitic on the other sanctions because firms have little incentive to abide by injunctions unless they fear penalties from doing so. If those penalties, in turn, are essentially the same as those for the underlying violation, it is not clear what an injunction adds with regard to deterring future violations. A higher sanction or use of more summary proceedings for repeat violations might be employed, but such could be done in any event. Injunctions do differ from other remedies in that, instead of attempting to induce compliance, they may implement it directly, such as by an agency regulating firms' prices going forward or restructuring an industry so as to make conditions no longer conducive to successful oligopolistic coordination. Such remedies, of course, are often quite costly in themselves.

Consistent with the view that injunctions are not important in the present setting, it is not apparent that they are heavily used.⁴⁴ Nevertheless, many commentators seem to be fixated on their centrality—often implicitly, in arguing against a direct prohibition on coordinated oligopolistic price elevation on the ground that it would be difficult to fashion an injunction commanding

44 For example, exploration of the web site and publications on workload and enforcement of the U.S. Department of Justice Antitrust Division reveals substantial information on fines and imprisonment, but not (that this author could readily locate) on injunctions. Examination of competition law treatises in the United States and other research does not reveal significant attention devoted to injunctions for price fixing. It appears that injunctions are used with regard to explicit, open practices, such as when an organization requires price fixing or employs particular facilitating devices. Most U.S. Department of Justice cases that involve price fixing are criminal, and injunctions do not appear to be employed.

compliance.⁴⁵ The objection is that abstention from coordinated behavior, when conditions are conducive to success, would somehow be unnatural or irrational and hence futile to require, and the asserted implication is that liability must be limited to the commission of specified acts, such as particular forms of interfirm communication.⁴⁶

45 See, e.g., Kaysen & Turner (1959, 143–144) (“By the very fact that we are dealing with practices—that is, conduct—the appropriate remedy is always the injunctive remedy: cessation of the practice.”); Turner (1962, 669); and Baker (1996, 47).

46 See, e.g., Areeda & Hovenkamp (2003, 150) (“In the preceding hypothetical, however, we cannot reasonably expect firm *L* to refrain from increasing its prices when it feels that the market would accept them, or the others to refrain from following.”); *id.* at 231 (cannot employ “a legal rule [that] tells the oligopolist to close its eyes to the immediate and direct market impact of its own output choices, as well as to the subsequent market impact of its rivals’ probable responses to its own output decision”); *id.* at 232 (cannot employ a rule that “tells each firm to ignore the profit-maximizing signals emitted by the market”); Chamberlin (1929, 65) (“Each is forced by the situation itself to take into account the policy of his rival in determining his own . . .”); Dabbah (2004, 268) (cannot make illegal firms’ behavior that constitutes profit maximization); Elhauge & Geradin (2007, 801–802) (referring to “the problem that firms [in] oligopolistic markets cannot avoid knowing their prices are interdependent when they set their prices, so that it would be hard to define any prohibition in a way that tells firms how to behave”); *id.* at 835 (“If so, how could one define the offense in a way that oligopolists could avoid behaving illegally? Is it practicable to ask them to ignore the reality of their price interdependence when making their pricing decisions?”); Hovenkamp (2005, 128) (objecting to the condemnation of interdependent behavior standing alone because “implicit in condemnation of any practice under the antitrust laws is that the defendant was obliged to behave in some other way than it did”); Monti (2001, 145) (arguing that remedies are infeasible as long as there exists a rational economic explanation for the oligopolists’ behavior); Scherer & Ross (1990, 342) (“How should oligopolists change their behavior so as to avoid breaking the law? Must they begin ignoring their interdependence in pricing decisions, when to do so would be irrational?”); Stroux (2004, 114) (“Imposing competitors to disregard their rival[s]’ behaviour would obviously be nonsense, as it would require them to behave irrationally.”); Turner (1962, 666) (“Particularly is this so when the behavior involved, setting the ‘profit-maximizing’ price in light of all market facts, is not only legally acceptable but vitally necessary to make competitive markets function as they are supposed to function.”); Ivo Van Bael & Jean-Francois Bellis (2005, 51) (“[S]ometimes it is only rational commercial behaviour which makes competitors align their conduct. In such a case undertakings should not be punished for doing what makes sense commercially.”); and White House Task Force on Antitrust Policy (1968, 5) (“antitrust law . . . cannot order the several firms to ignore each other’s existence”). Interestingly, most of these references postdate Posner’s work that is cited in the following note, yet they seem to ignore its basic point with regard to the deterrent effect of sanctions altering what firms would find rational to do.

In addition to the fact that sanctions change firms’ incentives, the sometimes-expressed position that interdependence is inevitable that is reflected in the aforementioned views—see also, e.g., Kaysen & Turner (1959, 27) (suggesting that recognized interdependence is “extremely likely” when concentration is even moderately high)—is not empirically well grounded. Evidence on successful prosecutions, see sources cited *supra* note 20, reveals the use of explicit and sometimes highly elaborate direct communications even in very concentrated industries, and industrial organization research for decades on the structure-conduct-performance paradigm implies that successful coordination often fails even in highly concentrated industries. See, e.g., Perloff, Karp, & Golan (2007, 33–34) and Schmalensee (1989, 971, 976, 988; 1987b). Moreover, most commentators who have offered an opinion on the subject (overlapping in part with those cited just above in this note) assert that coordination is actually quite difficult in the absence of explicit communications. See *infra* subsection 3.4.4 and note 76. Accordingly, both the empirical basis and apparent consensus behind the premise for this common argument seem on reflection to be lacking.

This complaint, however, is quite puzzling since the response is both obvious and has been offered before. The point is simple: what is natural and rational depends on whether sanctions are imposed. It may be natural and rational for a hungry shopper to steal an apple, for a youth with public artistic impulses to create graffiti, or for a greedy manager to embezzle funds—that is, if such were legal. However, if these activities are illegal and subject to heavy sanctions, engaging in them becomes irrational. Oligopolistic price coordination is no different. If there are no sanctions, a firm may find it attractive to follow a leader's price increase and to avoid undercutting the industry's supracompetitive price because of the allure of sustained supracompetitive profits. But if such acts are associated with sufficiently high penalties, then the firm would find it irrational to follow the leader's price increase and profitable to undercut elevated prices.⁴⁷ Therefore, the question of whether a command to refrain from coordinated oligopolistic pricing—whether issued through a particular injunction or by the law more generally—will succeed depends on the adequacy of expected sanctions. This adequacy, in turn, depends on detection and on the magnitude of sanctions. It does not depend on firms' managers possessing some mystical ability to engage in a form of reasoning heretofore unknown to humankind.

In addition to the choice of sanctions, a complete analysis of optimal enforcement must consider additional issues (*see, e.g.*, Polinsky and Shavell 2007), but they are set aside for present purposes. Among them are whether private suits should be used instead of or in addition to public enforcement, how private enforcement should be operated (permission of class actions, allocation of attorneys' fees, determination of who should be permitted to sue), how obligations for fines or damages should be allocated among the firms (particularly if the defendants do not constitute the entire market or if some are judgment proof), the strategic use of leniency policy, deciding which firms should be liable at all, and liability for attempts.

3. COMMUNICATIONS-BASED PROHIBITION

Part 2 sketches how to construct a regime to address coordinated oligopolistic price elevation. The means of detection and determination of sanctions follow

47 See Posner (1969, 1592 n.80) ("All I am arguing is that a deliberate restriction of output by competitors is conduct that rational men can avoid—and will avoid if it is made sufficiently costly to them to engage in it."); Posner (2001, 97–98) ("Tacit collusion is not an unconscious state. . . . The threat of a damages judgment for supracompetitive pricing will influence their pricing decisions; what would be irrational would be for the oligopolists to ignore such a threat."); and Posner & Easterbrook (1981, 333).

from the statement of the social problem to be addressed rather than from statutory interpretation or some other external command. The problem is a challenging one because it is difficult to detect successful coordination with sufficient frequency to achieve substantial deterrence while limiting false positives so as to contain the cost of chilling desirable behavior. There are a variety of techniques available, and it seems best to employ them in combination, with the relative importance depending on the strength of particular evidence and other circumstances of a given case or type of case.

The approach that emerges appears to be quite different from what most commentators believe that the law does and should require. That alternative, often characterized as involving a requirement of express or explicit agreement, tends to focus on a subset⁴⁸ of interfirm communications in determining liability.⁴⁹ A priori, this scheme seems unlikely to be best or even nearly so. Because it asks the wrong question—whether there exists an “agreement” rather than whether harm is likely—it misuses many types of evidence that are probative of the existence of the social harm in question, privileging a small portion that will often be unavailable and sometimes be unreliable. (Of course, when it is available and reliable, it receives heavy weight under the more catholic approach toward evidence presented in section 2.2.) Moreover, as will be developed in sections 3.4 and 3.5, it takes one category of evidence, on the conduciveness of conditions, and gives it negative weight—that is, opposite to what is implied by the objective of minimizing social harm—in the most consequential settings, if one accepts commonly advanced views (although these will be called into question). Accordingly, it is difficult to defend this approach

48 As this term in the text emphasizes, the commonly advanced approach, despite sometimes being discussed as if it involves a general prohibition on communications, is actually intended, even if implicitly, to be a selective one, as section 3.1 elaborates.

49 See, e.g., *In re High Fructose Corn Syrup Antitrust Litigation*, 295 F.3d 651, 654 (7th Cir. 2002) (opinion by Judge Posner) (“This statutory language [of Sherman Act Section 1] is broad enough . . . to encompass a purely tacit agreement to fix prices, that is, an agreement made without any actual communication among the parties to the agreement. . . . Nevertheless, it is generally believed . . . that an express, manifested agreement, and thus an agreement involving actual, verbalized communication, must be proved in order for a price-fixing conspiracy to be actionable under the Sherman Act.”). This manner of articulating the difference is also often mentioned by economists. See, e.g., Kühn (2001) (advocating that liability for oligopoly pricing should be based exclusively on the use of particular sorts of communication); Porter (2005, 147–148) (suggesting a distinction between direct and indirect communication); Werden (2004, 780) (there must be “some evidence of communications of some kind among the defendants through which an agreement could have been negotiated”); and Whinston (2006, 20) (referring to the law as prohibiting “talking” between firms); see also Motta (2007, 315). Similar depictions are offered with regard to law in the European Union. See, e.g., Black (2005, 341) (finding “high[ly] plausible . . . the view that [concerted practices under EU Article 101, formerly 81] involve communication”) and Jones (1993, 276).

even as a proxy technique that might be rationalized by the difficulty of detecting successful oligopolistic coordination.

Because the key elements of defining the social harm, detecting its presence, and determining the appropriate sanction have already been considered, there is a fundamental sense in which the analysis in part 2 is complete.⁵⁰ That is, there is no logical necessity to consider why one or another approach that fails to emerge from a systematic analysis of the problem might not be superior after all. This conclusion is all the more compelling with regard to commentators' favored alternative given the *prima facie* basis just offered for doubting its sensibility. Nevertheless, because a communications-based prohibition has nearly monopolized analysts' attention—the main exception being Posner's writing, the substance of which is largely ignored—it is fitting to undertake a thorough, side-by-side comparison.

Section 3.1 begins by defining this alternative. Although familiar in a rough sense, closer examination reveals it to be obscure, so some elaboration is required in order to render coherent any further assessment. Sections 3.2–3.6 mirror part 2 in considering social welfare, detection, and sanctions. Sections 3.4 and 3.5—on detection through the use of circumstantial evidence, and the resulting paradox of proof—are particularly notable because they elaborate an important and underappreciated manner in which the most commonly advanced method, if consistently pursued, is not merely incomplete and off target but potentially perverse. Rather than being simpler than the direct inquiry presented in section 2.2, the inference process is more complex and, of greater concern, is in conflict with the social objective of the legal prohibition.

3.1. Definition of the Approach⁵¹

Despite being long discussed and widely endorsed, the commonly favored view of the determinant of liability for oligopolistic coordination is difficult to articulate. Most commentators, regulators, and courts offer neither a canonical statement nor a series of well-constructed illustrations from which one can infer the scope of the contemplated prohibition. In writing on the U.S. rule, the modal statement by commentators is that there must be an express or explicit agreement—although some reject this formulation and indicate that the breadth is at least somewhat greater, perhaps including tacit agreements,

50 An important exception involves institutional issues, many of which might best be addressed differently under different approaches to detection and the definition of the rule for liability. See Kaplow (2011a, sec. V.A).

51 This section draws heavily on Kaplow (2011c), especially parts I and II; repeated references thereto are omitted.

whatever they might be.⁵² This choice of language as well as the archetypical example of secret meetings to discuss future prices suggests an emphasis on particular sorts of interfirm communication.

A close consensus on the criterion for liability may not exist. If there is one, this author has been unable to discern its content.⁵³ Moreover, it is important to explore the approach broadly, not dwelling on particulars, since we wish to know whether strengths and shortcomings are generic rather than possibly idiosyncratic to one or another formulation. Accordingly, it is helpful to state the communications-based prohibition in an abstract and general fashion.

Let us say that the set *X* is comprised of only (but all of) those communications (or other acts⁵⁴) that are deemed to constitute agreements, or conspiracies, or concerted actions, and thus to give rise to liability. All other communications are deemed to fall in the permitted set *X'*. That is, any acts (or clusters of acts) that are deemed legally sufficient for liability are elements of set *X*, and all others, set *X'*.⁵⁵ Communications in set *X* might be limited by mode: face-to-face meetings, letters, phone calls, and emails count, whereas hand or smoke signals—deemed to be in the set *X'*—do not. Or they might be limited by content: future prices may be a forbidden subject whereas present prices would be permissible; assurances might be prohibited whereas declarations of intentions or predictions would be allowed. Or they might be limited by the setting: statements in smoke-filled rooms could be prosecuted whereas

52 See, e.g., Hay (2006, 891–895) (discussing the lack of clarity in courts' use of the term tacit agreement) and Kovacic (1993, 14–21) (discussing courts' varying usage of the term tacit agreement). Articulating the current state of the law is challenging. As explained in detail in Kaplow (2011c, part III), some authoritative pronouncements fairly clearly embrace all interdependent behavior in the prohibition; others just as clearly rule it out; elaborations are intermittent and inconsistent; and key elements of practice (the use of plus factors, jury instructions on liability, and the rule for determining damages—and also the absence of many of the litigation behaviors implied by the paradox of proof) implicitly adopt the broad approach that requires only interdependence, even when employed by courts that purport to take a narrower view. EU law is less elaborated and similarly murky.

53 Usually none of the key terms are defined, and they are often elaborated in conflicting ways. The ambiguity is often so high that it is difficult, if not impossible, to compare views.

54 For convenience and because it seems in accordance with what most have in mind, this article will use the language of communications when referring to the sorts of acts that are in set *X*. The analysis that makes reference to this set, however, does not depend on this interpretation. This generalization is also useful in light of the difficulty even of defining communications for present purposes, on which see Kaplow (2011c, sec. II.B).

55 Because this approach essentially targets particular acts, those in set *X*, that have a tendency to facilitate successful oligopolistic coordination, while not targeting successful price elevation itself, one might describe the approach as one that renders price fixing, standing alone, per se legal (that is, legal in and of itself) and deems illegal only certain types of attempts or the use of certain means of accomplishing it, which amounts to deeming illegal certain facilitating practices. Viewed this way, there arises the further doctrinal question whether there must exist an agreement (or conspiracy, concerted action, or whatever) concerning the use of acts in *X*. See Kaplow (2011c, sec. II.B.2).

public announcements would be tolerated; permissible methods of price announcements may be more circumscribed when there are a few large buyers that occasionally place large orders than when there are significant numbers of small, non-repeat purchasers. Or these and further dimensions might be combined.

The central reasons for focusing on some delimited set of communications rather than on commonly used terms like agreement, conspiracy, and concerted action are that the latter are vague, their scope is contested, and they are most plausibly understood as close synonyms for interdependence, as explored at length in the predecessor to this article (Kaplow 2011c).⁵⁶ The last observation means that the prohibition would be of the sort examined in part 2, which most commentators reject and which would render further comparison unnecessary. To illustrate the predicament, consider a simple example in which two gasoline stations on adjacent corners in a remote location successfully coordinate their behavior to charge a significantly elevated price. If they achieve this objective through sequential price jousting involving responsive posting of price signs, most would deem there to be no agreement, but if they accomplish it through a secret discussion having precisely the same sequence and content, all would deem there to exist an express agreement. Suppose, indeed, that identical information is conveyed, identical states of mind are produced (achieving a so-called “meeting of the minds”), and identical actions result, having identical consequences. The sharply and unambiguously opposite legal outcomes must be due to the nature of the interfirm communications, not the substance of the resulting consensus between the firms.⁵⁷

Note further that what matters is the character rather than the existence of communications. After all, in the example just given, communication is equally present in both cases. Firms generally set prices, issue public statements, and engage in trade association activity that involves communication, much of which is not and would not sensibly be prohibited. In any event, the present formulation of the sets X and X' , with the generality it encompasses, does capture much of what seems to be envisioned by those who would reach different outcomes in the hypothesized example and, notably, who would describe

56 For example, agreement or conspiracy is commonly defined as involving a meeting of the minds or common understanding, which concepts clearly cover plain interdependence that coordinates on an elevated price. As mentioned, many focus on communications as the target of the prohibition, yet such are involved in myriad settings in which no agreement is said to exist. One might try to distinguish signals from language, but the existence of rich sign languages (among other points) renders this demarcation untenable.

57 This point can also be seen by comparing secret price-fixing discussions with seriatim press conference that convey the same messages and achieve the same results, the former being regarded as the archetypal express agreement and the latter as permissible.

or advocate that the legal prohibition be significantly more limited than one on any oligopolistic coordination that succeeds in elevating price.

Before proceeding with the analysis in sections 3.2–3.6, it is worth contemplating further the contours of this sort of communications-based prohibition. It is natural to inquire into the grounds for singling out some types of communication while excluding others. Even though most commentators seem to favor a rule that does just this, essentially no attention has been devoted to the question of how one would decide which sorts of communication (or acts more generally) should be deemed to be in the set X and which in X' .⁵⁸ The most plausible approach would involve balancing the benefits against the costs of including particular acts in the prohibited set. However, the benefits would presumably be those of better deterring coordinated oligopolistic price elevation and the costs those of greater chilling of desirable behavior. Hence, consistently pursuing this route replicates the direct approach elaborated in part 2. This point merely restates what is mentioned at the outset of this part: the suggestion that, once the analysis in part 2 has been performed, the task is complete. The result would not be sets X and X' of the sort commentators generally envision. Rather, it would be just the direct approach itself.

Suppose that one sticks with a prohibition that is limited to the types of communication sometimes described as involving express or explicit agreement. A further difficulty in any communications-based prohibition is posed by the flexibility and substitutability of modes of communication. If one prohibits talking, individuals can write. If writing is unavailable, there are hand signals, even full sign languages. Firms competing in an industry may know each other well and interact over extended periods of time, allowing for the development of subtle, yet effective means of communication. They have strong incentives to find some way to communicate and may do so if any channels are left open.

To help understand this challenge, consider the problem of functional equivalents. One option is to limit the triggering category X to a prespecified list of modes, content, or various combinations, meaning that there will inevitably exist fairly close functional equivalents, to many elements in X , that do not give rise to liability. This approach invites circumvention. On the other

58 This gap is remarkable. It may be partially explained by the earlier observation that commentators have not done much even to define the approach that they claim is embodied in existing law and/or should be. In attempting to state a sharper definition, the question of why some acts are included in the prohibition and others excluded from it would become more salient.

There has also been only a modest amount of analysis and little highly probative empirical evidence (that is, in realistic, relevant settings) on the role of communications in facilitating oligopolistic coordination, and relatedly on the implicit assumption that communications are more useful than actions (such as price moves) in sending credible messages (contrary to the familiar maxim that “actions speak louder than words”). For further discussion, see Kaplow (2011c, secs. II.B.3, IV.B.2).

hand, if functional equivalents—both existing ones and ones that may be developed over time—are included in set X, then there is no real limitation. After all, the function in question is to communicate sufficiently well to enable successful interdependent behavior. Thus, when effective interdependence is present, the function has been served and liability would be triggered. A functional approach therefore implies no limit to the inclusion of all successful interdependent oligopolistic behavior, which is contrary to the intention of those who would impose a communications requirement. Put another way, it is difficult to conceive of commentators' communications-based approach to liability except as one entailing a substantial degree of formalism in defining the sets X and X'.

3.2. Social Problem

The key point about the relationship between a selective communications prohibition and social welfare has been stated in the introduction to this part: the connection between the rule and purpose is indirect, the rule focusing on a subset of means that are related to adverse effects on social welfare rather than seeking directly to ascertain whether behavior is socially detrimental. Questions addressed in section 3.1 concerning the difficulty of articulating this approach—defining which acts are in X rather than X'—reinforce the point that the method is significantly detached from the social objective. Indeed, the more one attempts to define which acts are in prohibited set X with the objective of the prohibition in mind, such as by employing a cost-benefit test or by including functional equivalents, the more the approach tends to dissolve into the direct inquiry pursued in part 2.

More broadly, as mentioned, a large portion of probative evidence would not seem to be directly relevant under the communications-focused prohibition. As will be explored in sections 3.4 and 3.5, such evidence is still be considered, but it is ultimately legally relevant for making inferences about the likelihood that acts in set X were employed—which is only indirectly relevant to whether liability would be socially desirable—rather than making inferences about whether undesirable behavior has occurred. Moreover, this will be so even when the latter inference is more straightforward and reliable than the former inference.

Although this mismatch between approach and objective is fairly plain, it is worth revisiting the social welfare consequences of oligopolistic coordination to sharpen our understanding of the gap. Successful interdependence generates losses in static (mainly allocative) efficiency and dynamic efficiency (particularly involving excessive entry)—or, under a common view, losses in consumer surplus. These adverse effects are caused by price elevation. On their face, they

seem to be neither more nor less severe when the means by which such elevation is coordinated happen to involve one or another type of communication or other facilitating practice.

Subsection 2.1.2 did identify some possible benefits of oligopolistic price elevation, such as gains due to product variety that is usefully increased by the prospect of supracompetitive pricing. These benefits, however, depend on certain industry traits—notably, the nature of consumers' demands for different products—but not directly on whether communications (or other acts) in some set X need to be or are in fact employed in elevating price.

If one considers the matter further, one can imagine that there might sometimes be a connection. Notably, when product differentiation is particularly important, coordination is much more difficult, so one might suspect that it would only be feasible (if it is at all) when frequent, explicit communications are employed. In this instance, the need for such communications is associated with the net social benefit of limiting price elevation being below average, so perhaps a more permissive approach could be justified in such circumstances. This implication runs directly counter to that of the communications-based prohibition, which confines liability to situations involving highly explicit communications.⁵⁹ This particular observation is speculative and possibly of limited importance, and there may well be other settings in which the need to use acts in some set X in achieving price elevation is instead positively correlated with the likelihood or magnitude of social harm. Nevertheless, the relationship seems in most instances to be attenuated and not of sufficient strength to warrant an across-the-board, exclusive focus on the use of acts in set X.

Social welfare analysis is also importantly concerned with chilling effects. Perhaps substituting a selective communications prohibition has advantages in this regard. One possible influence on chilling costs depends on how high the burden of proof is set. One might suppose that it would be set quite high. In that case, there would be few false positives—and significantly diluted deterrence as well. Of course, one might set the proof burden high under the direct approach, similarly reducing false positives. Indeed, the prior discussion in subsection 2.1.3 explains that it probably makes sense to elevate proof burdens when key evidence is associated with significant chilling costs.

The more relevant question is whether placing greater (exclusive) reliance on evidence of the use of practices in X would better distinguish true coordinated price elevation from actual competition than would the alternative of considering all the relevant evidence, focusing on how it illuminates this distinction

⁵⁹ This limitation arises both *de jure*, by the nature of the prohibition on its face, and *de facto*, in that credible evidence of use of such communications is more likely to appear when they are more frequent and elaborate.

rather than a qualitatively different one, and giving the greatest weight to the types that are most reliable in a given context. The question largely answers itself. This subject should be kept in mind in the analysis that follows, particularly the discussion in sections 3.3 through 3.5 on detection and in section 4.1, which introduces a more radical approach that eschews any use of circumstantial evidence.

A further point suggesting an *a priori* basis for skepticism about the communications-based prohibition concerns the relationship between social welfare and the likelihood of finding liability by type of industry. Chilling effects are most likely, and the benefits of deterrence are least important, in industries that appear to be fairly competitive. However, because these industries tend to be less conducive to collusion, they are the ones—under the (possibly mistaken) views of many, as discussed in sections 3.4 and 3.5—in which the use of acts in X and thus a finding of liability are relatively more likely under the communications-based prohibition than under the direct approach. Therefore, upon further analysis, a communications-based prohibition seems to be even more poorly matched with the social objective than might first appear.

The comparison of approaches is also helpfully viewed from another, complementary perspective—one focused on the use of evidence and the setting of the burden of proof. As a general proposition, it would seem that a concern for chilling effects is best addressed through a combination of utilizing as much probative evidence as possible in a manner that reflects complementarities and risks of error and of setting an appropriately high proof requirement. Precisely these considerations dominate the analysis in part 2, particularly section 2.2. The communications-based approach reflects a qualitatively different strategy: instead of elevating the burden of proof directly, it changes what it is that must be proved. In doing so, the relevance and weight of evidence is determined by considerations that have a looser relationship to the social objective, and arguably a strong negative connection in important instances. This fundamental strategic difference explains why commentators' preferred method is likely to constitute an inferior, perhaps significantly inferior, alternative.

3.3. Detection: Internal and Other Direct Evidence of Prohibited Communications

Detection of prohibited communications—determining whether firms used any acts in X or confined themselves solely to acts in X'—is difficult for the usual reason that firms try to hide actions that are deemed to be illegal and subject to significant sanctions. In this respect, the challenge is similar to that in section 2.2, addressed to the detection of successful oligopolistic coordination. It is important to keep in mind, however, that what we are seeking to detect here is qualitatively different from what it was in the prior discussion. Even when

considering the same sort of evidence, we are not using it to answer the same question.

To begin, consider the relevance under a communications-based prohibition of whether successful oligopolistic coordination has taken place. Under one approach, liability would arise only when acts in X were employed and they succeeded in achieving price elevation. In that case, one would have to undertake the analysis in section 2.2 as well as that necessary to detect the use of prohibited communications. Under another approach, demonstration of the use of acts in X would be sufficient to establish liability. This regime would differ in that it would in essence punish attempts (of a certain type) independent of success.⁶⁰ Making the use of prohibited communications sufficient for liability may also be justified on the ground that direct proof of successful coordination is difficult while success may be inferred from the attempt, presuming that firms would not undertake the effort and risk liability unless they thought success was likely.

Most commentary does not mention this distinction, much less explore which approach is thought to be embodied in existing law (if either is) or constitutes better policy. In any event, as discussed in section 2.3, a sensible sanctioning regime tends to base penalties in significant part on the extent of success, in which case the magnitude of coordinated oligopolistic price elevation, if any, will be relevant even under the view that attempts are enough to establish liability. Moreover, as will become apparent, inferring from circumstantial evidence whether acts in X were used involves a preliminary determination of the existence of successful coordination. For these reasons, the analysis in this part will ordinarily assume that proof of success is required, one way or another, although the discussion will bear on failed attempts and on those whose success cannot be ascertained. Even so, the focus of this section and those that follow will be on how evidence affects the inference that acts in X were employed—rather than solely acts in X'—for the problem of detection of success was examined in section 2.2.

This section begins the analysis of detection of prohibited communications by considering internal and other direct evidence of proscribed communications. Sections 3.4 and 3.5 analyze market-based evidence of prohibited communications. This latter inquiry is extensive because the process of inference can be complicated in ways that are not well appreciated, and the implications for a communications-based prohibition are far-reaching.

If secret meetings among competitors in hotel rooms to discuss future pricing are prohibited, direct evidence might consist of a recording of a meeting,

60 On the punishment of attempts in the price-fixing context, see Kaplow (2011a, sec. V.B).

a witness to its occurrence, or an internal document referring to the event. Such proof is often referred to as “smoking-gun” evidence, and, when it is available, unambiguous, and reliable, the inquiry may well be complete. Due to firms’ interests in maintaining secrecy, however, this type of proof often will not be available. Additionally, internal evidence may be ambiguous and some, notably testimony of informants who may be disgruntled employees, may be unreliable. Also, such evidence may implicate particular individuals, but to establish the firm’s culpability it may also be necessary to consider the individuals’ authority in the firm, who else knew about their activities, and whether the firm likely acted in light of the communications.

As suggested by the prior discussion in subsection 2.2.2, internal evidence should also be considered much more broadly—and it ordinarily is in cases that lack smoking-gun evidence. Thus, all manner of internal evidence—planning documents, emails, data-gathering activities—may provide a basis for inferring whether acts in *X* were employed. In addition to direct, detailed references, there may also be indirect indicators, such as internal exchanges that may not have been possible in the absence of the prohibited interfirm communications. For example, there may be expression of knowledge about other firms’ future behavior, or discussions of strategies that would not make sense without such specific understandings. As mentioned previously, firms may attempt to avoid creating internal evidence or to distort its appearance, but it is unclear the extent to which such efforts can succeed.

With much internal evidence, the inference process is far more treacherous than is usually recognized. The reason is that the communications-based prohibition is premised on the view that interdependence may be possible without having to rely on acts in *X* and that such is legal. Keep in mind that the set *X'* is large, and it specifically includes many means by which firms might communicate with each other. To illustrate the difficulty, suppose that internal documents make explicit, repeated references to the firms having an agreement or engaging in concerted action. While these may be taken as admissions of liability, such would be a mistake under the proffered, restrictive view of the law. As noted, these terms are readily defined as interdependence and do not in ordinary usage convey information about the means by which such agreement or concert was achieved. Sometimes they do: a seller of a house saying that the property is under an agreement clearly conveys a conventional understanding of an explicit, typically legally binding, contract. But outside such contexts, a broader set of implications is possible. Accordingly, even direct statements of the existence of an agreement may in themselves do little to indicate whether such agreement was reached using at least one act in *X* or only acts in *X'*.

More often, one might see internal evidence of the existence of an understanding in an industry, that a firm does not want to behave aggressively

because such action would upset rivals, and so forth. Such evidence is likewise silent on how any such understanding came about.⁶¹ Therefore, much internal evidence may be highly probative of the existence of oligopolistic coordination (the question in section 2.2), but not, standing alone, very probative of whether communications in the prohibited set X were employed. Note importantly that, in such cases, proof of successful interdependence would be easier and more reliable than proof of the use of prohibited communications.⁶²

To overcome this obstacle, it is necessary to make the further inference that the agreement or understanding that is demonstrated through internal evidence could not, under the circumstances, have come about unless the firms engaged in prohibited communications. In this respect, once one moves past crisp, smoking-gun internal evidence (e.g., a document referring to the meeting at the Sands Hotel on April 7, 2010 at which a firm's vice president discussed and agreed with counterparties Smith from Rival 1 and Jones from Rival 2 to raise prices on May 1, 2010, from 100 to 120), one begins to enter the territory where one is also employing other evidence in attempting to make the requisite inference. This supplemental evidence may also be internal in nature. Often, however, internal evidence (plus any direct proof of prohibited interfirm communications themselves) will not by itself be sufficiently powerful. In such cases, one must rely in varying degrees on market-based evidence.⁶³

3.4. Detection: Market-Based Evidence of Prohibited Communications

3.4.1. Circumstantial Evidence

It is a basic maxim of competition law and of conspiracy law more generally that circumstantial evidence may be employed.⁶⁴ Moreover, it is widely

61 This point does not seem to be well appreciated. See Kaplow (2011c, 747–748 n.158).

62 It is as if the social objective is concerned with the presence of hay, but the rule forces us to wade through a substantial haystack in order to ascertain whether or not some needles might be hidden inside.

63 Note additionally that even the purely internal and other direct evidence—when not in itself conclusive—suffers from the problems raised by the paradox of proof that are explored in the sections to follow: if one accepts that the relationship is like that depicted, say, in Figures 1–3, the result is to exonerate defendants in cases posing the greatest social danger and the least risk of chilling effects, as discussed in section 3.5.

64 On U.S. antitrust law, see, for example, ABA (2007, 5–6 & n.29) (“Conspiracies can be proven either by direct or circumstantial evidence. . . . [C]ourts traditionally recognized that ‘[o]nly rarely will there be direct evidence of an express agreement’ in conspiracy cases. . . . Circumstantial evidence as to this element of the offense is . . . not only admissible, but often dispositive.” (quoting *Local Union No. 189, Amalgamated Meat Cutters v. Jewel Tea Co.*, 381 U.S. 676, 720 (1965), and for the latter proposition, citing, inter alia, *Monsanto Co. v. Spray-Rite Service Corp.*, 465 U.S. 752 (1984))), Areeda and Hovenkamp (2003, 2), and Blair (2008, 3) (who, in introducing a symposium on *Bell Atlantic Corp. v. Twombly*, 550 U.S. 544 (2007), mentions that “plaintiffs may, of course, rely upon circumstantial evidence”). On conspiracy law more generally, see, for example, LaFave (2003, 267)

accepted that, once acts are made illegal and subject to heavy sanctions, they will be driven underground and accompanied by concealment, so circumstantial evidence may need to be employed if adequate deterrence is to be achieved. This proposition is especially applicable in the present setting since the prohibition under discussion is limited to particular communications or other identified acts, those deemed to be in the set X . The relevant contrast is with a prohibition aimed at successful oligopolistic coordination, where the act in question, coordinated price elevation, is more public—even though, as section 2.2 indicates, it is nevertheless often difficult to detect.

Making inferences about the use of prohibited communications from circumstantial, market-based evidence is likely to be difficult. Some acts in X may be virtually indistinguishable from some acts in X' when viewed directly, clearly, and closely, particularly with regard to those near the boundary. How, then, are they to be distinguished when viewed indirectly, obscurely, and from afar? One must search for symptoms, but such will often be noisy signals. Even highly explicit and uninhibited communications can fail: some firms may not agree, and, even when they all think that they have agreed, cheating may nevertheless break out quickly. And communications limited to ordinary pricing behavior in the marketplace might succeed. It is usually believed that explicit, frequent, and direct communications make success more likely. The hypothesized relationship is probabilistic. On average, the difference in the likelihood of success when using certain elements of X rather than certain others of X' (those near the boundaries) will be small. Oligopolistic coordination may be quite unlikely, but nevertheless more likely when using acts in X . Or it might be highly likely, but even more so when using acts in X . Furthermore, as stated, these relationships are only true on average; in any given case, there might be no discernable effect (and sometimes the effect may even be the reverse). Hence, the routine inference problem is daunting.

Do note, however, that this problem will not be present in all cases. For example, liability may be found if observed behavior could only have arisen through the use of prohibited means in X that have no resemblance to any of those in X' . Unfortunately, such clarity may rarely be possible.⁶⁵ More broadly,

“It is thus well established that the prosecution may ‘rely on inferences drawn from the course of conduct of the alleged conspirators.’” (quoting *Interstate Circuit v. United States*, 306 U.S. 208, 221 (1939), an antitrust case, on this general principle of the criminal law of conspiracy)).

65 Some commentators have identified circumstances applicable in special cases in which inferences about communication might be more straightforward. Perhaps most often mentioned is the submission of identical, non-round-number, secret bids for made-to-order items. See, e.g., Areeda & Hovenkamp (2003, 168, 243). All of the qualifications are necessary for the claimed inference, which shows how infrequently it would be available. Even such limited exceptions have further, unrecognized qualifications. Notably, it must also be true that public price announcements are deemed

in order to make the necessary inferences, one will have to view all the facts and circumstances bearing on conduciveness to successful coordination, supplement these findings with knowledge of which means of coordination are associated with what probabilities of success in the identified type of setting, and also consider the degree of success that has occurred. How all of this may be done is considered next.

3.4.2. *Paradox of Proof: Overview*

The ideas underlying inferences about the use of prohibited communications from market-based evidence are as follows. First, one considers whether successful oligopolistic coordination is likely to be taking place. If it is very unlikely, then an inference that it is being attempted through improper means—that acts in X are being employed—is unwarranted, which is to say that the probability is

illegal (which most offering this sort of illustration do not believe to be the case). To see why, suppose that the use of prohibited communications might otherwise be inferred from the posited identity of secret bids at a price of, say, 3.518 per unit. To break the inference of secret advance communications, all that is required is for one clever firm to announce publicly that it intends to bid (or is thinking of bidding) 3.518, at which point a follow-the-price-leader explanation is a competing inference. Of course, the firms may well have met secretly to agree on the 3.518 price; but as long as, after the meeting, some firm makes the public announcement—and this too can be planned at the meeting—the ability to infer from the pricing coincidence that such a meeting must have occurred would be disrupted. (The fact that we never seem to observe this circumvention strategy is suggestive of what firms and their lawyers implicitly believe the law to be. See Kaplow (2011c, sec. III.E).)

This example also has another serious deficiency: if the firms are secretly meeting to set prices, why should they all submit identical bids that will appear suspicious? See Posner (2001, 87). Instead, they could—and, it appears, do—choose a low bidder in advance and arrange for others to submit plausible-looking-but-higher bids. See, e.g., Porter (2005, 156) (who, in reviewing a study of auctions for oil and gas leases, suggests that bids may have been submitted in such a fashion as “to create the appearance of competition”); Porter & Zona (1993) (finding that non-winning cartel members submitted phony higher bids in highway construction auctions); and Department of Justice (2010, 2) (“Complementary bidding schemes are the most frequently occurring forms of bid rigging, and they defraud purchasers by creating the appearance of competition to conceal secretly inflated prices.”). Winners might be chosen randomly by the colluding firms. However, more often, particular rotations would be more efficient and thus profitable for the conspirators. For example, it would be better to allocate particular bids to firms that have more excess capacity at the moment or are more favorably located. And there is evidence that bidders that meet secretly to plan bids do rotate in this fashion rather than submit identical bids and leave it to the buyer to choose the winner. See Comanor & Schankerman (1976) (arguing that bid rotation is more likely, especially when there are smaller numbers of firms, and reporting that, in prosecuted bidding cartels, the substantial majority of cases involving eight or fewer firms employed bid rotation) and Cook (1963, 68) (“In fact, if I were asked, I would certainly bet that most conspiracies involving public tenders are conspiracies to rotate the low bid—and the business. That is, the bidding firms have agreed among themselves who should get the business, and they purposely let that company be low.”); see also Davis & Wilson (2002, 64–72) (examining the effects of allowing communications on pricing in experimental sealed-bid auctions both when firms’ costs are fixed and when they depend on output committed as a result of winning previous auctions). Indeed, it has been argued that the high prevalence of identical secret bids in government contracting is evidence of less explicit cartel arrangements. See McAfee & McMillan (1992, 584).

low.⁶⁶ Since cases that involve no apparent coordination are ubiquitous—all the more so if deterrence is reasonably successful—the probability is presumably too low to justify a positive inference.

Now suppose that successful oligopolistic coordination is established with some requisite likelihood. Under the more direct analysis considered in section 2.2, the inquiry would be complete. However, under a communications-based prohibition, we must ask a second question: whether, under the circumstances of the market in question, such success is sufficiently unlikely in the absence of the use of prohibited communications (acts in X) to warrant an inference that such acts were used. This inference process is obviously more involved than that required under a prohibition on successful oligopolistic coordination since an additional inquiry is necessary. Moreover, as will be seen, this second question is more difficult to answer than is the first.

Before proceeding, the reader should be warned that implications of the inference process under examination are peculiar in that they differ in substantial, unexpected ways from prior understanding and sometimes seem greatly at odds with how parties and adjudicators view the law. The discussion here does not, however, purport to describe the law in action, but rather to draw out previously unexamined logical implications of conventional views. Conflicts that emerge bear on whether a communications-based prohibition makes sense and also on whether common understandings of what the law currently requires are sound.

Let us begin by stating the basic logic that gives rise to the paradox of proof (see, e.g., Fraas & Greer 1977, 29–30; Baker 1993, 185–194). In any case under consideration, there will exist various evidence about the extent of oligopolistic coordination that is occurring or has occurred. This factor and some others will be taken as given for the moment but explored further in subsections 3.4.3 and 3.4.5. The present focus is on the degree to which the market at hand is conducive to successful coordinated oligopoly pricing.⁶⁷

Start at the extreme end of the spectrum at which circumstances render successful coordination impossible regardless of the means employed. In that case, there is a strong inference that prohibited communications were not used.

66 It is not zero because the evidence on successful coordination may be mistaken (false negatives) and because attempts may fail.

67 For most of this section, the emphasis is on factors that influence the ease of coordination, but not all factors that are favorable to oligopolistic price elevation have this feature. For example, if the product in question is one for which there is greater market power—that is, the profit-maximizing monopoly price is higher—it may not be any easier for oligopolists to succeed in maintaining a supracompetitive price, although their motivation to do so would be greater. Hence, as market power rises, throughout the range it might be more rather than less likely that they would employ acts in X that facilitate coordination.

While parties might misperceive the situation and the evidence of infeasibility might be mistaken, suppose that neither possibility is very likely.⁶⁸

Next, consider moving from this pole to examine cases in which successful coordination, although still not easy, is increasingly plausible. Along the way, the likelihood that prohibited communications, that is, acts in category X, were employed rises. Firms would not risk such action, the direct detection of which would give rise to liability (including possible criminal sanctions), unless there was at least some likelihood of success. In this range, that likelihood is taken to be increasing, so the plausibility of the use of the necessary actions to achieve success is similarly increasing.

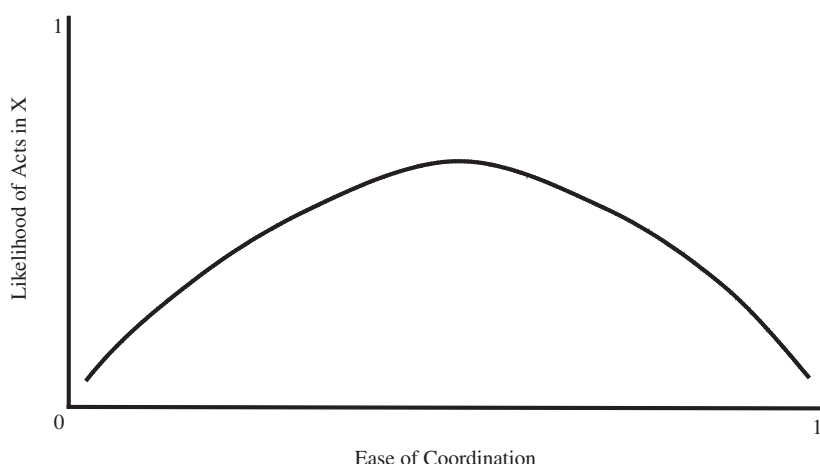
In contrast, consider the opposite end of the spectrum, settings in which successful coordination is so easy that firms can accomplish it almost automatically. In that case, it is thought to be unlikely that acts in X were employed. Why would firms use such means, with a risk of detection and possibly severe sanctions, when by assumption they can accomplish the same results confining their activities to those in X?⁶⁹

If one considers cases just short of this pole, the likelihood of the use of prohibited communications would plausibly be higher. Coordinated oligopoly pricing is taken to be sufficiently easy that there is a substantial chance of success, but as this chance becomes smaller, the effectiveness of coordination may well be enhanced by employing a broader set of tools, making it ever more likely that the firms resorted to some acts in X.

Combining the foregoing cases, it seems that the likelihood of the use of prohibited communications (or other prohibited behavior), defined as the use of acts in set X, at first rises and then ultimately falls as the degree to which the market is conducive to coordination increases. This characterization is represented in Figure 1. As one moves from the left to the right along the horizontal axis, the ease of coordination is taken to be increasing, starting with no possibility of coordination (denoted by 0) and ending at extremely easy coordination (denoted by 1). The vertical axis indicates the likelihood (probability, ranging from 0 to 1) of the use of prohibited communications, that is, of at least some act in X. As mentioned, it is supposed that initially, as the ease of coordination

68 Recall that the discussion takes as given evidence bearing on what pricing is actually present in the market. If there were, for example, strong evidence of consistent coordinated oligopoly pricing, then the likelihood of a mistake might actually be high. The alternative assumption is that strong evidence that such pricing is nearly impossible in the industry would cast doubt on the evidence that it is occurring. For the present, suppose we are not considering a case in which the evidence of successful oligopoly pricing is extremely strong. The interaction between these two dimensions is discussed in subsection 3.4.3.

69 Reservations are developed in subsection 3.4.5.

Figure 1. Ease of Coordination and Likelihood of Acts in X.

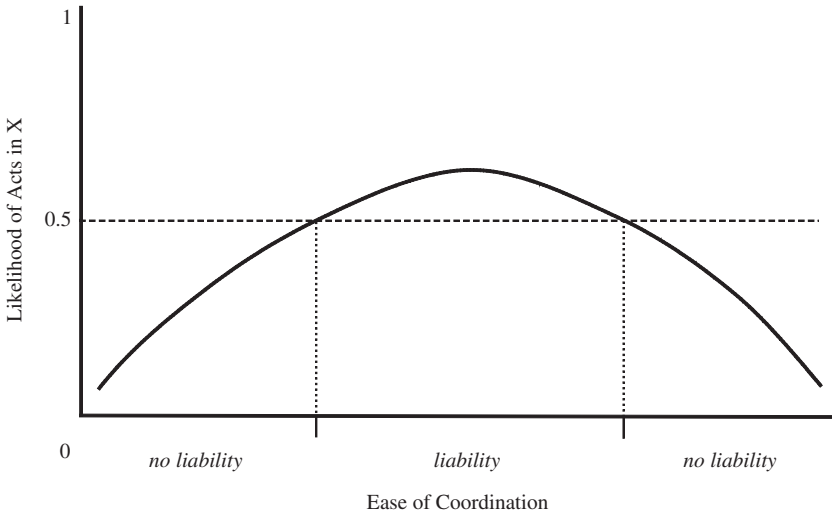
risks, starting from zero, the likelihood that some act in set X was employed rises, but, as the ease becomes sufficiently high, this likelihood falls. (The smooth, simple, and roughly symmetric character of the curve, as well as its height at various points, should be understood as an abstraction for illustrative purposes, on which more below.)

It is still necessary to translate this inference—from the ease of coordination to the likelihood of the use of acts in X—into a finding on liability. For concreteness, consider the proof standard of more likely than not,⁷⁰ in which case one simply needs to consider the portion of the curve that lies above 0.5 (50%), as depicted in Figure 2. (Such a liability threshold may well not be optimal, even conditional on the prohibition being defined in this manner, a doubt suggested by the analysis in section 2.1.3 on the optimal tradeoff of deterrence and chilling costs (*see* Kaplow 2012a, b).)

In this figure, we can see the immediate implications for liability. When the ease of coordination is low, there is no liability because the inferred likelihood of the use of prohibited communications is too low. When the ease of coordination becomes sufficiently high, in the middle region, there is liability. As the ease becomes higher still, the likelihood of the use of acts in X begins to fall and eventually becomes low enough that once again the likelihood is insufficient for liability.

70 For other proof standards, it is obvious how the following diagrams would be modified. *See infra* section 3.5. Relatedly, it will be discussed in subsection 3.4.3 that the curve in Figure 1 might in fact be higher or lower; the level depicted in this figure is essentially arbitrary. However, the qualitative statements presented throughout would be largely the same.

Figure 2. Ease of Coordination, Likelihood of Acts in X, and Liability.



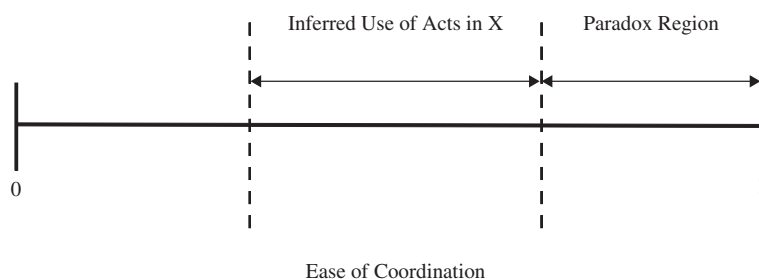
As a shorthand, consider the depiction in Figure 3. Here, attention is confined to the horizontal axis. The middle (liability) region is where the use of acts in X is inferred with the specified probability (taken here to be greater than 0.5). The intuition is that the ease of coordination is sufficiently high that attempts to collude are likely yet the degree is sufficiently low that collusion probably cannot succeed without using at least some act in X. The right (no liability) region is designated as the paradox region. There, the ease of coordination is so great that coordination is imagined to be readily accomplished without having to resort to any means in X.

The existence of this third, rightmost region may be regarded as paradoxical in a number of respects. Perhaps most obvious and important, this formulation makes clear that the legal requirement as conventionally understood carries the implication that there is liability in settings of moderate danger but exoneration in cases where the expected harm from anticompetitive behavior is at its greatest, a point that will be elaborated in section 3.5. Additional paradoxical features are developed in the next subsection.

3.4.3. Paradox of Proof: Implications for Adjudication

Using Figure 2 or Figure 3, it is easy to see the a priori ambiguity of presenting evidence that a given market poses a greater ease of coordination—for example, evidence of higher concentration, more nearly homogenous products, more readily observed prices, and larger numbers of smaller buyers placing more frequent orders. Further evidence that a market is more conducive to coordination than one might have believed in the absence of such evidence will be

Figure 3. Ease of Coordination and Paradox of Proof.



helpful to an enforcer (government agency or private plaintiff) toward the left of the horizontal axis. As the extent of such evidence mounts, however, the case eventually approaches the border with the right region, where the inference will revert to one of no use of acts in X and thus no liability, leading to a victory for the defendants. Therefore, whether evidence indicating a somewhat greater ease of coordination favors the enforcer or the defendants will depend on whether the other facts of the case suggest that it is nearer the left boundary or the right boundary of the middle, liability region.

This point has interesting implications for parties' litigation strategies. Starting with the enforcer at the point of composing a complaint, should it allege that conditions are highly conducive to coordination or the opposite? Regarding the ultimate outcome, convincing a decision-maker that conditions are somewhat more dangerous than what it otherwise would have concluded may be helpful (toward the left of the horizontal axis) or detrimental (toward the right). At the outset, it may not be obvious which situation will prevail, that is, which of the two boundary lines will, after elaborate proceedings, be seen by an adjudicator to be nearer to the case at hand. Additionally, this uncertainty may arise not only regarding projections about where on the spectrum the case falls but also regarding where the decision-maker will believe that the two boundaries are located (on which more in a moment). In any event, we should expect that a number of enforcers (especially in cases where the harm from oligopoly pricing and thus fines or damages would be the largest) would find it in their interest to allege that conditions are less conducive to coordination than meets the eye. Defendants' strategic interests are the reverse.

Of course, each side must not go too far. Enforcers win if the conclusion is that the ease of coordination is moderate, defendants if it is either high or low. In this state of affairs, we also should sometimes find parties arguing on the same side of the issue, for example, when the enforcer is concerned that the decision-maker may find coordination to be fairly difficult but the defendants think that coordination will be found to be fairly easy—or when the parties have different conjectures about where the decision-maker will place the two

boundaries. We might also see a party switching its side on these issues mid-stream, if it senses that the decision-maker is likely to land nearer to a boundary different from the one originally predicted. Overall, it is unusual that one party needs to show a moderate state of affairs and the other gains by demonstrating either extreme. (The defendants' argument in the alternative is of the form: "Coordination is quite difficult in our market; however, if it isn't quite difficult, it's actually very easy. In any event, it surely isn't anywhere in between.")

The matter also seems to make it difficult ever to grant a motion to dismiss, or for summary judgment, for a court would have to decide that the allegations, or also the undisputed facts, eliminate any serious question about the region in which the case falls. The problem is that, except in extreme cases, virtually any allegation or fact on these dimensions could cut in either direction on liability, depending on how other facts are weighed, something the court is not supposed to do at these preliminary stages.

Consider also the implications of the paradox of proof for the use of experts to demonstrate liability. A (perhaps the) central concern with the more direct approach developed in section 2.2 that focuses on detecting successful oligopolistic coordination is that it involves too complex an inquiry requiring extensive and subtle expert economic evidence that is subject to manipulation, particularly by financially motivated private parties. The approach to liability that depends on detection of prohibited communications through circumstantial evidence requires, as mentioned, answering this same question about the degree of successful coordination and also the one focused on here. Regarding the latter, the just-described inference process, as it would play out in adjudication, seems to raise even greater concerns.

The problem is actually worse than may first appear because of the need to determine the height⁷¹ and shape⁷² of the curve in Figures 1 and 2, which in turn determine the boundaries between no liability and liability in Figures 2 and 3. As a general matter, empirical evidence bearing on this relationship—which is contingent on both the strength of evidence that successful coordination is taking place and on the legal regime, that is, the definition of the set *X*—seems nonexistent and difficult to develop. Moreover, the nature of the relationship will depend on facts of the particular case (notably, particular features of

71 For example, if the curve in Figures 1 and 2 was substantially higher, the liability region would correspond to much of the horizontal axis. (In principle, it could reach either or both ends. Keep in mind that the curve is drawn taking as given evidence on the demonstrated likelihood and degree of success in oligopolistic coordination, which might be quite high.) Alternatively, the curve could be much lower, in which case the liability region would be thinner or even nonexistent. (Suppose that other evidence showed that the industry was probably exhibiting competitive pricing.)

72 Variations are presented in the two subsections that follow.

markets will affect the probability of success that may be achieved using particular combinations of acts and thus the likelihood that at least some act in X was employed). Accordingly, witnesses, expert or otherwise, will have to rely substantially on conjecture when supporting the requisite conclusions.

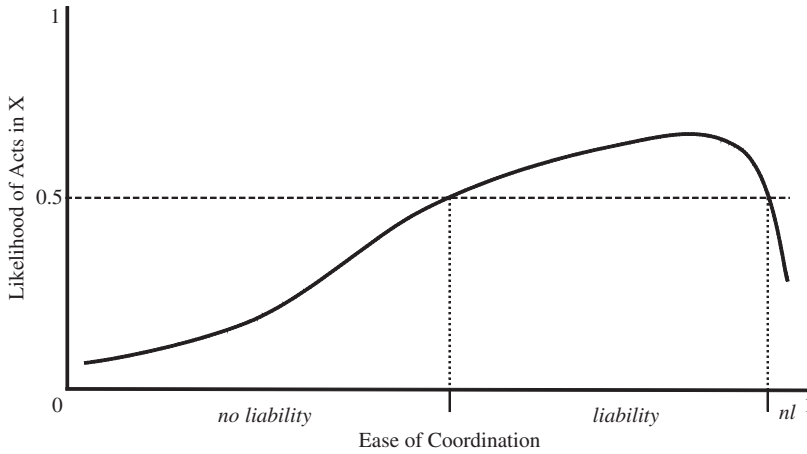
In comparing the direct approach examined in section 2.2 with the communications-based prohibition, we can ask which question economists—whether employed by an enforcement agency or supplying reports to and testifying before an adjudicator—can know more about: whether successful oligopolistic coordination is taking place or whether any observed pricing could only have come about using at least some act in a legally specified set X or may instead have come about using solely means in X'. However difficult is the first question in various settings, it is a subject of decades of theoretical and empirical research, whereas little is known about the second question, and even less with regard to any particular legal definition⁷³ of the sets X and X'.⁷⁴

Further confounding the inference process are a number of simplifying assumptions employed in subsection 3.4.2 that need to be relaxed to apply the framework in practice. One already alluded to is that a given set of diagrams is based on a particular finding relating to the degree of achieved success in oligopolistic coordination. Different degrees imply different curves and thus different regions for liability and no liability. And the degree of success will, of course, be contested. Thus, in principle, an adjudicator needs to know the proper curves for each possible finding that may be reached. Furthermore, evidence bearing on the ease of coordination may also bear on the extent to which success has been achieved, and vice versa. Indeed, each conclusion is related to the other: if success seems highly likely and substantial, it is implausible that the ease of coordination is very low; if the ease of coordination is low, then one would be more skeptical of evidence demonstrating success; and so forth. Viewed broadly, one can consider a mapping from all the evidence to a finding of liability, abstracting from separate, intermediate conclusions—a viewpoint applicable to the present inquiry into the likelihood of use of prohibited communications, acts in X. This perspective is reinforced by considering that there also may be some internal or other direct evidence of the use of prohibited communications that needs to be incorporated in reaching an ultimate decision on liability.

73 That is, any expert opinion on the likelihood of use of some acts in X given some set of facts will depend on precisely which acts or combinations of acts are deemed to be in set X versus in X', a matter that itself will be disputed.

74 To this we can add the query emphasized in section 3.2, above, and in section 3.5 to follow, concerning which question is most probative of whether the application of sanctions would enhance or reduce social welfare. On both counts—feasibility and normative relevance—the choice does not seem close.

Figure 4. Narrow Paradox Region.

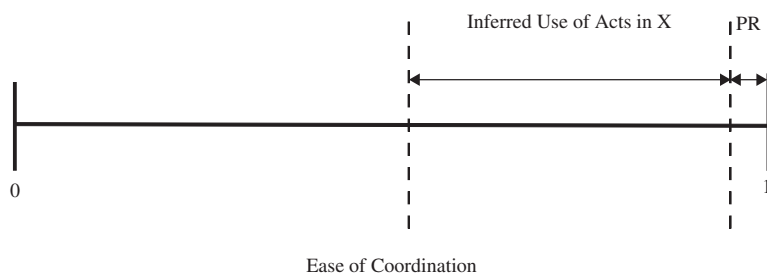


3.4.4. Breadth of Paradox Region

The preceding subsections consider for illustrative purposes a single, simple, symmetric relationship between the ease of coordination and the likelihood of use of prohibited communications. But the relationship could take other forms as well. Suppose that the relationship between the ease of coordination and the likelihood of use of acts in X is instead as depicted in Figure 4. Just as in Figures 1 and 2, the likelihood of use of acts in X first rises and then falls as the ease of coordination increases. The differences are that it takes a greater ease of coordination for the likelihood of the use of acts in X to exceed 50% and that this likelihood remains over 50% until the ease of coordination is quite high. The “nl” in the figure is an abbreviation for “no liability,” the shortened label reflecting that indeed the paradox region is quite narrow in this case, as shown in Figure 5.

This case is particularly interesting because, despite the existence of the paradox of proof, it looks almost like the more intuitive story (suggested by the analysis in subsection 2.2.1.2) in which there is no liability when the ease of coordination is low and liability when conduciveness is high. Except in or near a narrow, extreme region, in which conduciveness is very high, evidence of a greater ease of coordination favors (or at least does not seriously hurt) the enforcer and contrary evidence helps the defendants. If these figures depict the actual state of affairs (again, for a given legal standard, summarized by X and X', and given evidence on the extent of oligopolistic price coordination), there would be little difference in practice from a legal rule that deemed interdependent oligopolistic coordination sufficient for liability in those settings in

Figure 5. Narrow Paradox Region and Paradox of Proof.



which the ease of coordination was sufficiently high, specifically, beyond the left boundary in Figures 4 and 5.⁷⁵

Of course, there remains a difference between the postulated rule, under which oligopolistic coordination is sufficient for liability when the ease of coordination exceeds a threshold, and the rule embodied in X and X' that implicitly lies behind Figures 4 and 5: in those cases in which the danger is especially high (the worst cases), the latter rule would exonerate the defendants rather than finding them liable. A policy-based defense of the X/X' approach would have to be grounded in the desirability of allowing coordinated oligopoly pricing in such cases, a point that will be considered more broadly in the section 3.5. It would also have to justify expending the resources necessary to distinguish them from somewhat less extreme cases—keeping in mind that the availability of the defense, even if it is not often true, may motivate defendants to advance it often, both raising costs and producing some erroneous exonerations in the process.

It remains to ask whether the present depiction is plausible. Many analysts speak as though this case governs. In offering this characterization, writers do not speak directly in these terms, for the paradox of proof is only occasionally recognized and it is not analyzed in the manner done here. Instead, what one finds are strong statements regarding the difficulty of any real success at oligopolistic coordination in the absence of fairly explicit means of communication, that is, acts in the set X , even if that set is defined rather narrowly; a number of commentators deem instances of successful plain interdependence to be “rare.”⁷⁶ Given such beliefs, it follows that something like Figures 4 and 5

75 This depiction rests on the further assumption that most cases do not fall at the extreme right of the horizontal axis, which supposition is supported by the claims described below in the text.

76 See, e.g., Areeda & Hovenkamp (2003, 215) (“The results of perfect express collusion will rarely be achieved by mere interdependence without an express agreement standardizing some product or price terms or relationships.”); Hay (1981, 445) (suggesting that the dilemma regarding whether to attack all interdependent behavior or just express agreement may not be that practically important because pure oligopolistic interdependence is likely to be rare); Motta (2004, 141) (offering reasons

describes the regions of liability and the paradox. Although most who advance this view offer no empirical evidence in support,⁷⁷ the industry characteristics of cases subject to price-fixing prosecutions are consistent with this supposition: many involve small numbers of firms and homogeneous products, that is, settings in which the ease of coordination is high—and nevertheless secret interfirm communications were employed (*see, e.g.,* Hay & Kelley 1974, 26–27; Motta 2007, 318).

It seems surprising that so many commentators implicitly assert conditions implying a narrow paradox region since most also insist that the law does not and should not make successful interdependent oligopolistic coordination a basis for liability.⁷⁸ First, as noted, there is little difference in practice between

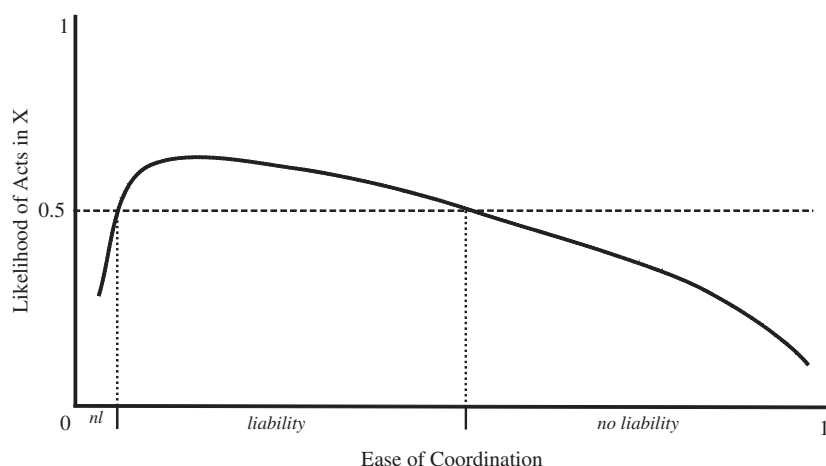
that it would be difficult for firms to coordinate successfully without talking to each other); Motta (2007, 317) (“it is far from clear that tacit collusion can be sustained over time without competitors talking to each other”); Neven (2001, 57) (“We review what the economic literature, and in particular the literature in experimental economics, has to say on the matter. We find few reasons to think that collusion is ‘easy’ to undertake without extensive ‘concertation’ between firms.”); Nye (1975, 209) (statement of then-Commissioner of the FTC) (“To put it more forcefully, I believe that pure ‘interdependence’ without some form of express collusion, however collateral, is a rare case, perhaps almost academic.”); Posner (1969, 1574) (“[I]t seems improbable that prices could long be maintained above cost in a market, even a highly oligopolistic one, without some explicit acts of communication and implementation. One can, to be sure, specify an extreme case in which such acts might be unnecessary. No more than three sellers selling a completely standardized product to a multitude of buyers (none large) should be able to maintain the joint maximizing price without explicit collusion. However, not many industries resemble this model.”); *id.* at 1575; Posner (1976, 904) (referring to the fact that “one can imagine a group of sellers able to collude without any overt contact or communication” but asserting that “[s]uch a case is probably rare”); Posner (2001, 66–69); Turner (1962, 665) (“some finite minimum of explicit communication, at some time, is involved” (quoting Kaysen (1951, 268–269))); *id.* at 672–673; Werden (2004, 762–763) (“it is far less clear that unspoken agreements are a significant phenomenon”; for a “considerable time . . . a widely held view . . . is that ‘coordination cannot be simply spontaneous’” (quoting Elzinga (1984, 25))); and Whinston (2006, 26) (“most economists are not bothered . . . perhaps because they believe (as I do) that direct communication (and especially face-to-face communication) often will matter for achieving cooperation”).

77 Neven (2001, 66) supports his claim by referring to experimental evidence on cooperation in games. There is, however, reason to be skeptical about the relevance of such evidence in the real-world contexts of present concern. *See, e.g.,* Whinston (2006, 24) (questioning “whether the results of these experiments, usually with college students as subjects, are indicative of the actual market behavior of businessmen and women”).

78 Posner (2001, 97–98), whose academic writing favors reaching successful interdependent oligopoly pricing, has a different take.

Anyway there probably are few cases of purely tacit collusion. What is being proposed is less the alteration of the substantive contours of the law than a change in evidentiary requirements to permit illegal price fixing to be found in circumstances in which an actual meeting of the minds on a noncompetitive price can be inferred even though explicit collusion cannot be proved. In most of these cases there will be explicit although well-concealed collusion that can certainly be deterred by threat of punishment.

Figure 6. Broad Paradox Region.

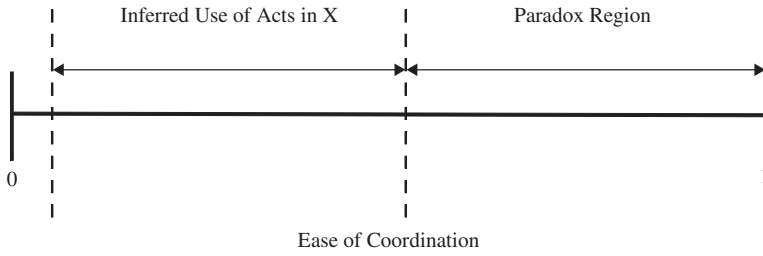


such a rule and one that carefully delineates a set X if one should virtually always infer the use of acts in X when successful coordination is demonstrated. Why is there so much insistence on one rule over another when they differ so infrequently? And why might it be thought that the more restrictive rule would filter out a substantial number of cases, especially at the pleading or summary judgment stages, as is often supposed? After all, if there is basis for supporting oligopolistic interdependence, then sufficiently great ease of coordination is powerful evidence favoring liability (even when defined as more limited than interdependence); only in exceptional cases would the ease of coordination be so great that liability would be negated. Finally, as noted just above, the little actual difference seems to favor the interdependence rule, for the divergence involves only the cases posing the most extreme danger.

Now suppose instead that the relationship between the ease of coordination and the likelihood of prohibited communications takes the form depicted in Figure 6. Once again, just as in Figures 1 and 2, the likelihood of the use of acts in X first rises and then falls as the ease of coordination increases. The difference now is that it takes little danger of coordination for the likelihood of prohibited communications to exceed 50% and this likelihood remains over 50% only until the ease of coordination is moderate. In this instance, the narrow “nl” region in the figure is at the left extreme. This scenario can be restated as in Figure 7.

This scenario contrasts sharply with that just discussed. Here, the paradox region is quite broad, so we would expect enforcers’ and defendants’ strategic situations often to be reversed, notably, in all cases in which the ease of

Figure 7. Broad Paradox Region and Paradox of Proof.



coordination was moderate to significant. Rare would be disputes where it was important for an enforcer to establish that the ease of coordination was greater and where defendants would wish to prove the opposite, for few cases would fall in or near the leftmost region of Figure 7.⁷⁹

The present setting, with a broad paradox region, is one in which the choice to define agreement more restrictively, in terms of sets X and X' (where, moreover, the set X is not too inclusive), has great consequences. Most commentators seem to believe that high stakes rest on whether agreement is thus defined, rather than being interpreted to encompass oligopolistic interdependence generally. Accordingly, one might expect to see frequent endorsements of factual propositions consistent with Figures 6 and 7, but (as noted just above) the opposite is the norm—that is, statements are suggestive of Figures 4 and 5, with a narrow paradox region. Furthermore, because belief that the debate over the rule of liability is important implies that the X/X' implementation involves a broad paradox region, one would expect great attention to be devoted to the sorts of reversed positionings of the parties described in subsection 3.4.3. That is, if one were to suppose that most seriously contested cases are not at the far left of the horizontal axis, then it would not merely be possible but typical for enforcers to argue from the outset (the filing of their complaint) that collusion was difficult and defendants that the industry was highly conducive to coordinated oligopolistic pricing. Yet this too does not appear to be so. In sum, it seems that much commentary on the subject, in failing to attend in a sustained manner to the paradox of proof, reflects a lack of appreciation of the implications of different choices of the legal rule.

⁷⁹ It is possible that most filed cases would be frivolous, in the sense that there was indeed virtually no prospect of collusion. In those cases, the left boundary in Figure 7 would be the operative one, and conventional positionings could be expected.

It is also useful to consider the differences between Figures 5 and 7—or 4 and 6. As drawn, the outcome in nearly all fact situations (for just about any degree of ease of coordination) is reversed in the two depictions. (The exception is at the extremes.) In other words, for a given legal rule in the form of X/X' , the circumstantial evidence could be completely clear (even stipulated by the parties) and yet, in most cases, opposite results would be possible. A case could be toward the left end of the spectrum (short of the extreme), and either the enforcer or the defendants might be entitled to win; likewise toward the right end. Who should win would depend on which of the two curves relating the ease of coordination and the likelihood of acts in X is the true one, reinforcing the importance of the analysis in subsection 3.4.3. Accordingly, it would be important for parties to present evidence on this question—although, as already noted, little is known. Furthermore, uncertainty about which situation prevails in a given case implies that there now is a double strategic paradox. The second is that an enforcer would wish to argue that Figures 4 and 5 are correct—if the case is likely to end up on the right half of the horizontal axis—but it would wish to argue that Figures 6 and 7 depict the truth—if the case is likely to end up on the left. And, again, the defendants' strategic stance would be the opposite.

3.4.5. Countervailing Effects, Dependence on Legal Rule, and Endogeneity of Firms' Behavior

There are a number of further complications in addition to those already explored. Initially, let us revisit the basic assumptions behind the paradox of proof diagrams, beginning with Figure 1, regarding the standard view that implies that the likelihood of use of prohibited communications falls with the ease of coordination once the ease becomes sufficiently high. This feature, which is necessary to generate a paradox region, can be questioned on two grounds.

First, it is supposed that helpful acts in X become less valuable when conditions are more conducive to coordination, at least once conduciveness exceeds low to moderate levels. This relationship, however, need not be true (ignoring potential liability). Simply put, the value of practices in X may well be increasing, not decreasing, as conditions improve, even at high levels of conduciveness. Certain sorts of communication may raise the expected average markup from, say, 10% to 15% when conditions are moderately conducive but from 30% to 50% when they are very conducive. For example, price wars are far more costly when prices are elevated substantially rather than modestly, so avoiding them is more valuable in the former case. Accordingly, even toward the right of the figures, incremental benefits from the use of acts in X may rise rather than fall as the ease of coordination increases—and, if and when they do fall, they may not decline until the ease of coordination is quite high and they may not fall as much as seems generally to be supposed.

Second, the paradox logic implicitly assumes that the risk of sanctions from using acts in X is constant along the horizontal axis. However, some factors that make coordination easier also make detection less likely. When there are fewer firms, coordination is simpler, *ceteris paribus*. But it is also true that explicit communications can be confined to fewer actors, making detection more difficult (see Hay & Kelley 1974, 20, 23–24 & n.15, 26–27; Masson & Reynolds 1978, 25–26). The probability of leaks and other means by which enforcers may detect the communications falls, making the use of such communications less dangerous. Similarly, if products are homogeneous or the environment is more stable, there is less need for complex and frequent negotiations, which likewise reduces the likelihood of detection. If the expected sanctions from direct detection consequently fall reasonably rapidly, the likelihood of the use of prohibited communications may be steadily rising as the ease of coordination increases. It might then be true that only permitted means are employed to attempt to achieve modest success when conditions are merely moderately conducive but prohibited means are employed when conditions are more conducive. Again, the result would be no paradox region at all.⁸⁰

For these reasons, suggestions by some commentators and courts that highly conducive conditions negate any inference that prohibited communications were employed might be wrong. Indeed, such conditions would not only fail to negate the inference, but they may instead strengthen it for most or even all of the right portion of the diagrams, perhaps producing the situation depicted in Figures 8 and 9.⁸¹ In such cases, there would be much less of a difference between liability triggered by mere interdependence and a more restrictive rule that limited liability to cases involving the use of prohibited communications (or other specified acts, defining the set X). Under the direct approach of part 2, one could choose a proof burden for conduciveness (taking as given the evidence on the existence of successful oligopolistic coordination) that mirrors the result produced under the communications-based prohibition. With only a single boundary between no liability and liability, the remaining question would be where optimally to set that boundary. This determination should be governed by the analysis in subsection 2.1.3 directed at maximizing social welfare, not by some calculus designed to meet a target likelihood that acts in some set X were employed.

80 A related point is that the absence of smoking-gun evidence not only fails to indicate in general that explicit, secret communications did not occur, but this lack of implication is especially notable in cases in which conditions are highly conducive.

81 More moderately, the points in this subsection might explain why there would be a narrow paradox region, as discussed in subsection 3.4.4 and depicted in Figures 4 and 5.

Figure 8. No Paradox Region.

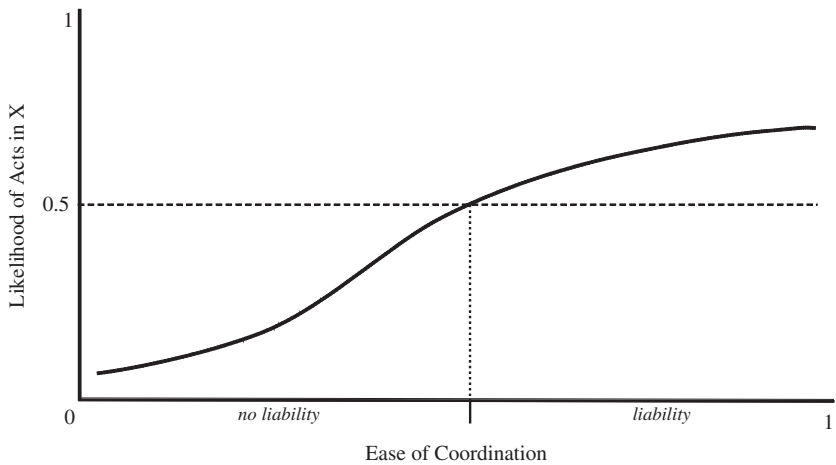
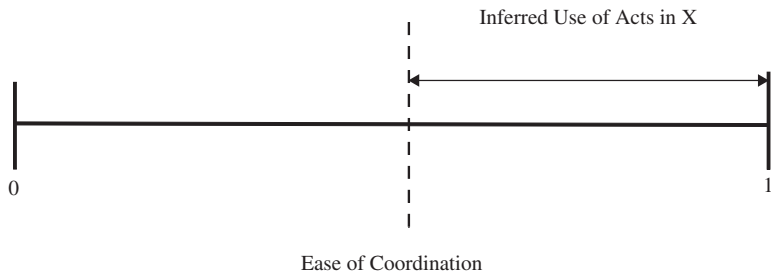


Figure 9. No Paradox of Proof.



The analysis throughout this section is also conditional on an assumed legal prohibition, that is, a stipulation of which acts are deemed to be in set X versus set X' . Obviously, if X is defined more narrowly, it will tend to be less likely that a prohibited act was employed. As the definition of X is broadened, the curves in the preceding figures will rise—the likelihood of some act in X having been employed will be greater for a given ease of coordination. Furthermore, it is apparent that a legal rule, and thus the corresponding sets X and X' , must be defined rather precisely. After all, the factfinder must infer, often from complex and remote circumstantial evidence, how likely it is that the defendants confined themselves entirely to acts in X' rather than having had to use at least one act in set X . Knowing, say, merely that X is narrower than interdependence alone (under which no particular act is required) tells us very little. Likewise, knowing only that X is broader than a very precisely defined sort of explicit

agreement (in which case only a few sorts of acts would be in X) provides negligible guidance. As mentioned, those purporting to describe or advance a legal prohibition in this form tend to offer neither an operational definition nor a sequence of concrete illustrations of prohibited and permitted acts, so it is quite difficult to know what set X is envisioned. Accordingly, it is mysterious how a judgment on the likelihood of the use of prohibited communications is supposed to be reached. When one adds that, under the paradox of proof, much evidence may cut either way—depending on other evidence (and on beliefs about the height and shape of the pertinent curve)—it is even harder to understand the bases for inference under this approach.

Closely related is the level of enforcement. In suggesting that the likelihood of the use of prohibited communications is low when the ease of coordination is particularly low or high, the reasoning relied on the point that firms would not risk sanctions by employing means that had little prospect of success or when success was likely without resorting to them. However, if expected sanctions were sufficiently low, firms would nevertheless employ illegal means in these contexts. Similarly, in the middle region, where the likelihood of use of acts in X is relatively high, the actual likelihood would be lower if expected sanctions were greater; after all, the main purpose of sanctions is to achieve deterrence.

In addition, in making inferences about whether defendant firms were likely to have employed acts in some given set X rather than confining themselves to practices in set X', further assumptions were made about firms' understanding of the legal system. However, these ignore an important interaction between how firms chose acts and the method of inference employed. Notably, if the determination of whether firms employed practices in X is based on inferences from circumstantial evidence, and if such inferences are less likely to be drawn when conditions are highly conducive to successful oligopolistic interdependence, then why wouldn't firms for this very reason become more rather than less inclined to undertake acts in X when the ease of coordination is high?

Such analysis is incomplete, in part because there will sometimes be another route to liability, using direct evidence. It is generally believed, however, that, conspirators' ability to conceal their activity renders deterrence insufficient unless circumstantial evidence is also considered. This desired supplemental deterrence is what can be undermined by the logical inference process. Thus, there appears to be a conundrum: If the use of prohibited communications is not inferred because it is reasoned that firms will refrain from acts in X in the stated circumstances, then there is no additional deterrence of such acts and they will tend to be used. If the use of acts in X is inferred in the circumstances, then there would be deterrence and the acts would tend not to be used. In either case, the inference would be incorrect.

The lesson is that all of the argument up until this point on the paradox of proof needs to be modified—and, unfortunately, made more complex in yet another manner—as a consequence of the endogeneity of firms' behavior. It is insufficient to reason that firms simply would or would not employ acts in X because of the expected consequences. As a result, the best inference as to the likelihood of the use of prohibited communications is at some intermediate level. Firms' decisions to employ any such acts will depend on how they expect their acts to influence factfinders' inferences regarding that likelihood. The more an act raises the expected likelihood of liability, the more it will be deterred, but such deterrence reduces the likelihood that it is logical for a factfinder to assess. Accordingly, an inference process aimed at determining whether the likelihood of the use of prohibited communications exceeds 50% (or any other target probability) will need to find the balancing point (equilibrium) between these forces.⁸² (As subsection 2.1.3 mentions, however, it is generally optimal to determine liability instead by balancing the social welfare benefits of deterrence against the costs of chilling desirable behavior, both ex ante considerations—and there is no reason to believe that such an approach will yield some stipulated ex post likelihood of undesirable behavior.⁸³)

Subsection 3.4.3 raises a number of complications concerning the inference process implicitly (and logically) entailed by a selective prohibition on communications (or other specified acts). The process is daunting due to the level of complexity and the lack of an empirical basis for making the requisite determinations. Considerations raised by the present subsection suggest that this preliminary assessment was overly optimistic. Taken together, the analysis in the whole of this section indicates that implementation of a communications-based prohibition is immensely more challenging than is appreciated by its advocates and also notably more so than entailed by the more direct approach considered in part 2, which itself is often quite difficult to apply.

On reflection, it is apparent that commentators have not systematically considered the implications of taking seriously the sort of legal proscription that they claim is reflected in existing law and constitutes sound policy. When this is done, the result has a surreal air. The implications of the prevailing notion seem

82 Rational adjudicators interacting with rational firms (who all know that others think in the same fashion) might be expected in equilibrium to employ mixed strategies: adjudicators might expect firms to employ acts in X with a certain probability, which gives rise to probabilistic liability, with firms employing acts in X that often. In reality, unobserved heterogeneity of firms might result in determinate strategies. (If the probability of detection or the level of sanctions is sufficiently low, it is instead an equilibrium for firms always to employ acts in X and for adjudicators always to infer that acts in X are employed.)

83 This point is true more generally when determining optimal proof burdens. See Kaplow (2012a, b).

quite fanciful, casting doubt on its descriptive validity (*see* Kaplow 2011c, part III). However, these difficulties are distinct from the normative problem that is the focus of this article: even if done well, the communications-based prohibition is likely to be clearly inferior in terms of its effects on social welfare.

3.5. Comparison of Approaches: Deterrence, Chilling Effects, and Proof Burdens

The most important question in this article concerns how the two approaches differ in terms of outcomes and resulting economic consequences. To compare the direct approach to liability developed in section 2.2 with the communications-based prohibition considered here, it is useful to specify the former in terms of the above diagrams, say, Figures 2 and 3. Suppose that the direct approach would find liability when the ease of coordination is at least at the left boundary of the liability region in those figures, or perhaps starting somewhat further to the right. Any outcome is *a priori* possible; selection of the optimal boundary point, which can be understood as the optimal proof burden on conduciveness taking as given other evidence, including that on successful coordination, would be determined as discussed in subsection 2.1.3, on which more in a moment.

Regarding deterrence, the communications-based prohibition would be weaker (even nonexistent, depending on how well firms can predict factfinding) in the right, paradox region. Because this region is where the social danger—the expected harm from coordinated price elevation—is the greatest, this disadvantage of basing liability on the presence of particular communications is substantial. In this regard, one should also keep in mind that, as explained in section 2.1.2, the net expected harm from oligopoly pricing rises disproportionately with the magnitude of price elevation; for example, when elevations are likely to be twice as great, social harm will be significantly more than double.⁸⁴ Likewise, under the communications-based prohibition, deterrence is maintained in the middle region, and (depending on how proof burdens are set under the direct approach) may be, in comparison, relatively strongest toward the left of that region, which is where expected harm per event is lowest.

Consideration of chilling costs also seems adverse to the communications-based prohibition if one considers the costs per unit of enforcement or deterrence. In an absolute sense, if liability is reduced—say, the prospect is the same in the middle region and negligible in the paradox region—total chilling costs will fall. But we should be concerned with the significance of chilling costs relative to deterrence benefits. Toward the middle of the diagrams, where

84 However, the loss in consumer surplus, although still increasing with the price elevation, does so at a decreasing rate. *See* Kaplow (2011b).

conduciveness is moderate (including cases toward the left of that region, where it is not very high), errors in identifying successful coordination are greatest, as explained in subsection 2.2.1.2. All else equal, it is less plausible that coordination is taking place when conditions are relatively uncondusive. Moreover, any success achieved is likely to be smaller in magnitude, and mistakes are more likely when elevations are small. Likewise, error seems less likely as one moves toward the right in the diagram and least likely in the paradox region. In sum, the communications-based prohibition aims at cases where collusion is not as easy, which will tend to include a relatively higher fraction of cases in which collusion is less likely to exist, so false positives that produce chilling costs will tend to be relatively more frequent. Furthermore, chilling costs are most serious in industries that are truly competitive, or very nearly so, which will be more likely to be true as one moves toward the left along the horizontal axis.

Combining these two observations suggests that a communications-based prohibition exonerates firms in cases in which the tradeoff of deterrence benefits and chilling costs is most favorable to liability and imposes sanctions when the tradeoff is relatively less favorable. To dramatize this point, compare this prohibition to a direct one that imposes a particular burden of proof on conduciveness: specifically, such that liability begins at the border between the middle and right regions, that is, at the left border of the paradox region. In other words, liability in the middle and right regions is the opposite of what would prevail under the communications-based prohibition. And, to ease the exposition, suppose that there are the same number of cases in the middle and right regions.⁸⁵ Compare Figure 3', slightly altered from the original Figure 3, with Figure 3'', reflecting the just-described modification.

This comparison provokes a simple question: all else equal, how do we feel about imposing liability just in cases that pose moderate danger (Figure 3')—the situation under the communications-based prohibition—or just in cases that pose high danger (Figure 3'')—as under this particular implementation of the direct approach of part 2? Since deterrence benefits are higher and chilling costs are lower, possibly to a great degree, under the direct approach, it would obviously be superior. That is, there exists a way to set the proof burden under

85 This assumption is unrealistic; it would only hold by chance. Moreover, the number of cases in each region depends on the liability rule, and what we care about for ex ante effects are the number of settings influenced rather than the number of cases. In any event, the discussion in the text is merely suggestive and one could modify the burden of proof to produce the desired comparison. For example, to get the same number of affected markets under each rule, the proof burden under the direct approach of part 2 might turn out to be further to the left. Then we would have an intermediate region—from that point on the horizontal axis up to the border with the paradox region—where the treatment under the two rules was the same, and the difference would be between the remaining regions, which by construction would affect the same number of settings.

Figure 3'. Ease of Coordination and Paradox of Proof (Modified).

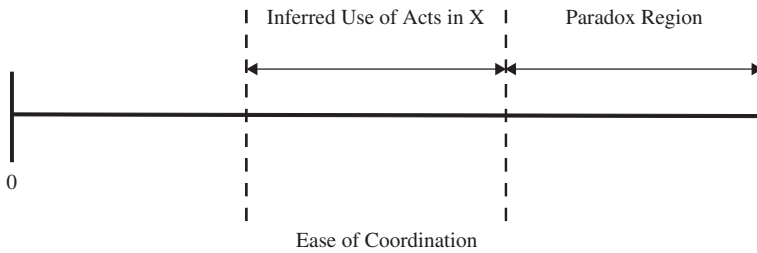
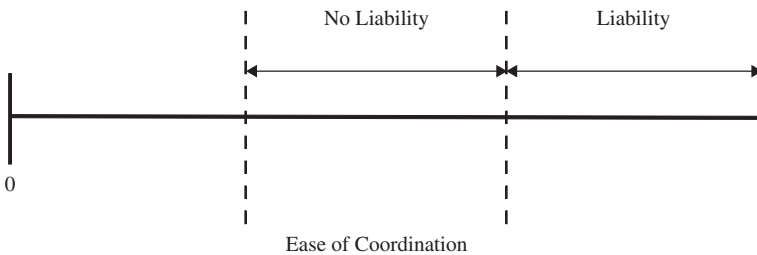


Figure 3''. Reversing Liability and Paradox (Rightmost No Liability) Regions.



the approach of part 2 such that it clearly, and perhaps overwhelmingly, dominates the communications-based prohibition.⁸⁶

More broadly, the concern with chilling costs due to the prospect of false positives is optimally confronted directly. Notably, it may make sense to increase the proof burden on whether successful oligopolistic coordination has taken place, rather than focusing only on conduciveness, which is but one consideration bearing on the central question of concern. Likewise, it may be appropriate to give less weight to particular types of evidence of more questionable reliability. In contrast, it is not optimal to exonerate defendants in a wide range of cases that pose the greatest danger and least chilling cost, including those in which the overall proof is strongest. Put another way, because a greater ease of coordination tends to be associated with both greater deterrence benefits and lower chilling costs—which is why it is identified in subsection 2.2.1.2 as a consideration favorable to liability—it is truly perverse to count it backwards in the strongest cases for liability.

The preceding point focuses on setting the proof burden under the direct approach. It is also revealing to consider adjusting the burden of proof with

⁸⁶ The proof burden that gives rise to this sharp comparison will, of course, be best only by chance. The optimal burden might be higher or lower; but, in any case, since it is optimal, the results would necessarily be superior to those under the burden stipulated here for ease of comparison.

regard to market-based inferences under the communications-based prohibition. To see the effects, recall the horizontal dotted line in Figure 2 that indicates where the probability of use of acts in X is 0.50, so that above that line such use is more likely than not. Raising the burden would involve setting a higher probability threshold and thus raising the horizontal line; conversely for a lower burden. The resulting tradeoff from either adjustment has a mixed quality. For example, raising the burden slightly may well be desirable with regard to the left boundary between liability and no liability because those cases in which liability is no longer deemed to exist pose a low danger of social harm and a significant possibility of chilling costs—that is, relative to the group of scenarios in which liability was initially assessed. However, the effect may be adverse at the right boundary, between the liability and paradox regions, because those now-exonerated cases involve relatively high danger and low chilling costs—compared to cases in the interior of the middle region, where liability is retained. That is, raising the proof burden under a communications-based prohibition exonerates both the weakest and strongest cases in terms of the social welfare consequences of liability.

This unfortunate—indeed, perverse—tradeoff is avoidable. Under the depicted circumstances, moving each boundary toward the right would be a better way to improve social welfare—supposing that liability is indeed desirable for some of the cases in the liability region. Doing so would entail a relatively higher burden toward the left of the horizontal axis—exonerating the weakest cases, as with a straightforward rise in the burden of proof—and a relatively lower burden toward the right—now resulting in broader liability rather than greater exoneration for the strongest cases.

Similar logic favors moving the right boundary all the way to the right—eliminating the paradox region—and setting the left boundary of the liability region wherever the tradeoff between deterrence benefits and chilling costs is optimized. But, of course, this realignment would abandon the communications-based prohibition embodied in the sets X and X' and replace it with the direct approach from part 2. This conclusion is hardly surprising; the reasoning merely recasts the logic from just above.

This point restates the theme with which this part opens, namely, that part 2 states the nature of the social problem and then considers how to go about detection and setting sanctions in an optimal fashion, that which maximizes social welfare. Since nothing like the communications-based approach emerged from that analysis, we knew from the outset that such a liability regime would not be optimal. Moreover, if we state some liability regime that differs from what is optimal and ask how it can be improved, we might expect that reshaping its boundaries to move it toward the previously identified optimal regime would make sense. The present analysis makes this

idea concrete by articulating the initial differences and explaining how the deterrence and chilling effects change as one liability regime is morphed into the other.

In this regard, it must be recalled from section 2.1.3 that the welfare-based approach does not define false positives and false negatives relative to some formal statement of a liability rule but rather in terms of how findings of liability and of no liability in various settings will influence firms' behavior, for better or worse, in terms of effects on social welfare. Commentators who favor a communications-based approach object to liability in the paradox region and thereby view liability in such cases as involving false positives. But such false positives are defined by reference to this formal standard for liability, whose desirability is in question. As discussed throughout this section, the prospect of liability in the paradox region tends to be especially valuable in terms of deterrence and less costly in terms of chilling effects than is liability in the middle region, where the posited approach would hold firms liable. The core "chilling effect" of liability in the paradox region consists of deterrence of coordinated price elevation, and it is deterrence in settings involving the greatest social harm. Hence, when the social objective motivating the legal prohibition is the focus of the analysis, the error tradeoff is assessed quite differently—and properly. This divergence between a rule based on the use of acts in a set *X* and one designed in light of the social objective is also what explains why the simple device of adjusting the proof burden upward or downward (raising or lowering the horizontal dotted line in Figure 2) is such a blunt—and odd—way to fine-tune a legal regime.

3.6. Sanctions

The subject of optimal sanctions is considered in section 2.3. Discussion here focuses on differences that arise when employing a communications-based prohibition rather than when basing liability on the presence of successful oligopolistic coordination.

First, as mentioned before, many (but not all) of the methods of demonstrating successful oligopolistic coordination also directly indicate the magnitude of price elevation and thus provide a basis for setting fines and damages. When liability is triggered by the use of some prohibited act rather than proven oligopolistic price elevation, additional inquiry for purposes of calibrating sanctions will often be necessary. One implication of this point is that, even in cases in which there is decisive internal or other direct evidence of prohibited communications, the expense and complexity involved with assessing market-based evidence will not be avoided.

Second, with a communications-based prohibition, there is an additional problem that concerns the pertinent baseline. As section 2.3 notes, the benchmark for setting fines and damages is (in principle) related to the extent of firms' overcharge. Under the direct approach, the overcharge is the difference between the actual price charged and a competitive price. However, if the violation is defined in terms of the use of prohibited communications, the use of acts in a set X , then it may seem logical for the benchmark not to be the competitive price but rather the possibly elevated price that the firms could have charged had they confined their behavior to acts in the set X' . After all, such price elevation is deemed to be legal under this regime, so the incremental harm attributable to the firms' illegal activity is limited to the magnitude of the further price increase made possible by using acts in X . It is not the entire gap between price and marginal cost at the competitive level of output.⁸⁷

Accordingly, a government intending to levy a fine that was a function of harm caused by the violation or a plaintiff hoping to recover damages would need to show not only that illegal methods were used but also how much of the price elevation could be attributed to the use of such means. In cases in which proof is by circumstantial evidence and the factfinder infers that the defendants' actions barely crossed the line demarcating liability, damages under this formulation could be quite small. But the implications of having to identify the increment above plain, interdependent oligopoly prices are not so limited. Even in a clear case of express conspiracy—smoke-filled rooms, taped conversations, and criminal convictions—it is possible that, but for the prohibited interactions, the firms would have elevated price above a competitive level, although perhaps not as high, for as long, or with as few price wars.

This additional requirement could prove quite challenging. First, one would have to determine which of the acts the defendants employed are in set X and which other acts they might have used in addition or could have used instead are in the set X' . Then, one would have to assess the degree of price elevation under both scenarios. That is, one would need to know the extent of successful price elevation in fact (even if this is not required for liability, established

87 Basing damages on this full gap seems more in accordance with current practice, thus constituting one of the respects in which existing law in action is more consistent with successful oligopolistic coordination being sufficient for liability. For elaboration on this logic and on the law in the United States, see Kaplow (2011c, sec. III.C.4). In brief, the general rule is that damages must be causally related to the actions that give rise to liability. See, e.g., Areeda et al. (2007, 332–333). However, this general principle does not seem to be followed (assuming the validity of commentators' characterizations of existing law on liability), with damages instead ordinarily reflecting the degree to which price is elevated above the competitive level. See *id.* at 377. Although this discrepancy has gone largely unnoticed, the point has been raised by the defendants' expert in at least one important case. See White (2001, 28).

perhaps by evidence of interfirm communications) as well as what it would have been if the defendants' had confined themselves to acts in X' . In determining the latter, there may be dispute over which acts, not necessarily ones actually used, would be legal (in addition to expected disputes over which acts actually found to have been employed were legal, even granting that some were illegal). And, of course, there would be significant disagreement about the extent to which prices might have been elevated if only such legal means were employed. Given how little is understood about this sort of question and the hypothetical nature of the inquiry, the need for this additional determination seems quite problematic.⁸⁸ Furthermore, even if this determination can be made well, deterrence of socially harmful price elevation is undermined thereby—although this problem is better attributed to the rule of liability that permits successful (even and perhaps especially very successful) coordinated price elevation as long as prohibited communications are not (and we can now add, need not be) employed.

The main point with regard to fines and damages is that further complication is introduced by the communications-based prohibition, in addition to the detection difficulties that are the focus of sections 3.4 and 3.5. The social consequence is a more costly and error-prone adjudication process in addition to one that is less effective in advancing social welfare even without regard to administrative concerns.

In contrast, the analysis of imprisonment of individuals, usually within a firm, for violation of a selective prohibition on communications is fairly similar to that under the more direct approach. One difference might concern the relative importance of imprisonment as a supplemental sanction. As previously explained, a key rationale involves remedying any deterrence deficit present when only fines (and private damages) are employed. Because firms may be sanctioned in a narrower class of circumstances when liability requires proof of the use of an act in set X , there may be less deterrence due to a lower probability of sanctions and thus arguably a reason to employ higher sanctions, possibly including imprisonment. However, what matters is whether firms know *ex ante* the situation they will be deemed to occupy *ex post*. Perhaps some firms that would have anticipated sanctions under a prohibition on successful oligopolistic coordination will expect sanctions under the narrower prohibition with a

88 Similar challenges would be posed in determining overcharges in standard ways, which often involve comparisons to other markets or to behavior before or after the period of illegal activity. Notably, there may be disputes about whether acts in X or solely acts in X' were employed in those comparison markets or time periods. To be sure, challenges arise with these comparisons under both approaches because it may be necessary to determine the extent of interdependence (however accomplished) in the comparison markets.

positive—because they are using acts in X, or might be found to have done so—but lower probability. Others, those well into the paradox region, may expect to be found liable with a low probability under a communications-based prohibition, so raising sanctions may not have much deterrent effect on them. A possibly countervailing consideration concerns chilling costs. Although there may be negligible cost from chilling acts in X, the communications-based prohibition disproportionately targets settings with a lower danger of coordination and a correspondingly greater likelihood of striking competitive behavior (as explained in section 3.5), so lower sanctions than otherwise might be optimal.

Another difference concerns the evidence required to identify the particular individuals who are to be subject to imprisonment (or personal fines). First, this process may seem easier under the communications-based prohibition since the prohibited class of acts might be defined more clearly—although the opposite is also possible, as suggested by the longstanding lack of a clear definition of what such a prohibition entails. Second, some evidence that might be clear under a broader prohibition may be ambiguous under a narrower one. Except for individuals caught on tape, it may be difficult to prove that others, such as their supervisors or other individuals who may have acted on the achieved interfirm understanding, knew that the resulting meeting of the minds was obtained using means in the set X rather than solely through acts in X'. It seems that it would be especially difficult to identify particular individuals with the requisite knowledge in a case proved primarily by circumstantial evidence. Of course, under the broader prohibition on successful oligopolistic coordination, the proof burden for sanctions against individuals may need to be high, and it might primarily be met when there is powerful internal evidence or proof of specific, explicit interfirm communications linked to the individual defendants in question.

The analysis of injunctions in section 2.3 establishes a number of propositions that are also pertinent to a communications-based prohibition. First, it is not obvious that injunctions are important (or that they are frequently used). This point is especially significant with regard to deterrence but also has implications for subsequent compliance because fines and imprisonment could likewise be used without any need for an injunction, and, when there is an injunction, it is still necessary to use other sanctions to ensure that firms have an incentive to comply with it.

This latter point is not obviously much different when the prohibition applies to particular communications rather than to successful oligopolistic coordination. Suppose that firms are caught having used some acts in X and that an injunction proscribes future use of those acts (or perhaps a broader set than those used, such as all acts in X). Proving the original violation may have been difficult, as suggested by the analysis in sections 3.3 and 3.4. If it was

proved by catching individuals on tape, one might have to catch them again, in the future, to prove that they violated the injunction. Even if violations continue, it might be difficult to detect them because of inherent obstacles as well as the fact that firms will be extremely cautious, having been caught once. Internal monitoring might be imposed, but this supplement would aid future deterrence regardless of the type of sanction employed. Alternatively, future violations might be established by circumstantial evidence. In either case, just as before, it would be necessary to replicate the detection and sanctioning process, which is little different from what must be done in the absence of an injunction.⁸⁹

Consider the case in which the available proof of a continuing violation consists primarily of ongoing successful oligopolistic coordination. Should it be inferred that some acts in X were employed? If such was previously inferred from this type of evidence, on the grounds that the conduciveness of conditions placed the case in the liability region rather than the paradox region, the same inference might be drawn again. At that point, there may be little difference (going forward) between a communications-based prohibition and a broader proscription on coordinated pricing. This result should not be surprising since the two rules tend to reach similar outcomes in such settings.

Alternatively, defendant firms might argue that, even if it was appropriate to infer their use of prohibited communications in the past, such should not be inferred in the future because, having learned well about each others' predilections and methods of signaling, elaborate coordination is now possible without resort to acts in X. Nevertheless, one might associate the future behavior with the prior use of acts in X and thus deem it to be attributable to illegal activity under the communications-based prohibition, which would constitute a further move in the direction of a broader prohibition on successful oligopolistic coordination. As discussed in section 2.3, many commentators would be bothered by such a regime since they seem not to comprehend how rational agents could desist from profitable behavior (price elevation). But it was explained there how, once sanctions are contemplated, the behavior is no longer profitable (if expected sanctions are sufficiently high), so self-interest would produce the desired effect.

89 It might seem appealing to switch the burden of proof to the defendants once they have been demonstrated to have committed a violation, but it is not clear how this would function. After all, there is no direct way to demonstrate the relevant negative—that the firms are no longer using any acts in X—except by a failure to prove the positive directly or through circumstantial evidence.

4. ADDITIONAL CONSIDERATIONS

4.1. Alternative: Disallowance of Circumstantial Evidence

Much of the difficulty with a communications-based prohibition arises in connection with the use of circumstantial evidence, which gives rise to the paradox of proof. Inferences are complex, costly, and error-prone, and, even when the process works well, the results are problematic in that cases posing the greatest danger lead to no sanctions whereas liability tends to be concentrated on cases involving less social harm but greater chilling costs. Accordingly, it may make sense to consider an alternative formulation of the communications-based prohibition that seeks to avoid some of these problems by disallowing the use of circumstantial evidence—that is, a regime that still requires proof of the use of an act in some specified set *X* (rather than solely the use of acts in *X'*) but imposes the further requirement that such use be demonstrated entirely through the use of direct evidence. On one hand, such an approach is radical in a formal, legal sense. As mentioned in subsection 3.4.1, it is a central tenet of competition law and conspiracy law more broadly that circumstantial evidence may be employed. Moreover, use of such evidence is generally thought to be particularly important in realms such as the present one in which parties hide activities that give rise to liability. On the other hand, criminal price-fixing prosecutions (that occur mainly in the United States), which target individuals for incarceration, seem to require direct, explicit, even smoking-gun evidence of prohibited communications, presumably due in significant part to high proof burdens and perhaps also the exercise of prosecutorial discretion. In any event, this more restrictive approach to liability (for all sanctions) deserves some attention.⁹⁰

First, chilling effects tend to be reduced when a legal prohibition is narrowed. Reducing the use of less reliable evidence is particularly likely to be helpful in this regard. This factor—along with savings in administrative costs—is the chief motivation for considering this more restrictive alternative. This benefit, however, may require that the range of permissible evidence be sharply limited. False positives and concomitant chilling effects may still arise due to the ambiguity and unreliability of direct evidence that is not of a smoking-gun variety.

90 A number of additional issues should be noted: there is no clear boundary between circumstantial and direct evidence (even a tape recording is only direct evidence of talking, not of whether the speakers are serious, have authority, or took any actions). As mentioned below, much internal evidence will be ambiguous. Informants may be untrustworthy. Market-based evidence may still be required to calibrate sanctions. And such a highly circumscribed approach does not obviously solve the dilemma addressed in section 4.4 regarding the screening of cases—namely, that if such evidence is required up front, deterrence may suffer greatly, but if it is not, then few cases may be eliminated at early stages.

For example, an enforcer may identify various internal fragments, such as from emails, that can be interpreted to demonstrate illegality, even though defendants offer contrary interpretations. Furthermore, exclusion of circumstantial, market-based evidence may raise the pressure to accept dubious but direct proof and will also sometimes undermine defendants' attempts to demonstrate that they actually behaved competitively. Nevertheless, if a sufficiently tough limitation is imposed, false positives would be less likely, possibly by a substantial margin compared to the unrestricted communications-based prohibition analyzed in part 3.

Second, disallowance of circumstantial evidence would presumably reduce deterrence because liability could no longer be successfully established in certain settings. One could attempt to remedy the shortfall through higher sanctions.⁹¹ There are, of course, limits to this stratagem, including its selectivity: across-the-board increases in penalties may be adequate or excessive in some settings (where firms believe that discovery of their use of acts in the set X is reasonably likely) but remain insufficient in those in which firms know in advance that their actions are unlikely to give rise to liability under the circumscribed rule.

Successful oligopolistic coordination may not involve any acts in X, and, for that which does, it is unknown what portion generates the sort of direct evidence that a highly restricted prohibition would credit. For the former group, we already know that a communications-based prohibition, requiring proof of acts in X, exonerates firms in cases involving the greatest social danger and relatively low chilling concerns. Among those cases that remain, there is the further point raised in subsection 3.4.5 that detection tends to be most difficult when there are few firms and the coordination problem is easiest, that is, the cases involving relatively large harm within the group of cases that actually involve acts in X. Finally, empirical evidence does not give us confidence that existing deterrence is adequate, both because there appears to be significant oligopolistic price elevation in many parts of the economy and because the very number and magnitude of price elevations in prosecuted cases suggests a deterrence shortfall, indicating that the probability of detection may be quite low (see Kaplow 2011a, sec. II.D).

For this circumscribed communications-based prohibition to have any chance of being on a par with the more direct approach of part 2, it would

91 In addition, one might increase enforcement intensity, such as by undertaking more undercover operations, employing a lower threshold for wiretapping, pursuing weaker leads, and so forth. The extent to which this alternative is feasible or desirable depends on the aggressiveness of current enforcement, among other things.

have to be that market-based and other circumstantial evidence—when used correctly—adds little to the ability to achieve deterrence of coordinated oligopolistic price elevation⁹² while notably raising the likelihood of false positives that generate chilling effects and also augmenting administrative costs. If this is true, note that it is the radical reduction in admissible evidence—rather than the fact that one is generally looking for the use of acts in X rather than evidence of oligopolistic coordination—that generates the alternative rule’s possible advantages. It follows, therefore, that if such limited evidence sufficed to achieve substantial deterrence, other reliable evidence rarely existed, and the chilling costs of allowing further channels of proof (relaxing the proof burden) were great, essentially the same result would be obtained under the direct approach. The main difference is that sharp, direct evidence of the socially undesirable phenomenon—successful coordinated price elevation—rather than powerful evidence of a symptom that is not a necessary condition to harm—the use of an act in some set X—would be the trigger for liability. The difference would be small if virtually any evidence unambiguous enough to be admitted involved both.⁹³

4.2. Implications for Other Competition Law Rules

It is understood that different aspects of competition policy are interrelated; what is optimal in one area depends in part on how other rules are set. The most important application of this principle concerns the oligopoly problem. It is

92 See also *infra* section 4.3 (on the possible erosion of smoking-gun evidence due to innovation in electronic communications).

93 If a highly circumscribed approach to addressing price elevation by direct competitors was optimal, in significant part due to doubts about factfinders’ abilities, it would seem that the law on monopolization, vertical practices, and mergers—areas that rely much more heavily on complex economic evidence and circumstantial evidence more broadly—should be extremely restrictive if not a nullity. After all, the social dangers in these other areas are often more ambiguous. Some commentators skeptical of the value of competition regimes endorse such a view. See, e.g., Bork (1978, 405–407); Crandall & Winston (2003, 24) (“Until economists have hard evidence that the current antitrust statutes and the institutions that administer them are generating social benefits, the Federal Trade Commission and the Department of Justice should focus on the most significant and egregious violations, such as blatant price fixing and merger-to-monopoly and treat most other apparent threats to competition with benign neglect.”); and Easterbrook (1984, 3) (“But suits against mergers more often than not have attacked combinations that increased efficiency, and the dissolution of mergers has led to higher prices in the product market. There are good theoretical reasons to believe that the costs of other enforcement efforts have exceeded the benefits.”). How restrictive a skeptic should be will depend on institutional considerations, such as whether adjudication is by expert tribunal, generalist judge, or lay jury and on whether prosecutorial discretion can be relied on to help moderate possible excesses of broader prohibitions. (For the latter to be effective, the prosecutor may need a monopoly on enforcement authority, for if private parties or numerous enforcers in other jurisdictions may each independently decide whether to bring a case, then it is the behavior of the most aggressive enforcer that will determine the actual scope of application.)

generally believed that the difficulty of directly attacking interdependent oligopoly pricing warrants a tougher policy toward horizontal mergers in order to render markets structurally less conducive to coordinated price elevation (*see, e.g.*, Baker 1993, 199–207; Brock 2006, 280; Kaysen & Turner 1959, 44–45, 127–141).⁹⁴ Of course, greater strictness comes at a cost in terms of forgone gains in productive efficiency and inhibition of the market for corporate control that provides important discipline for managers.

Conversely, to the extent that competition law is more aggressive against successful oligopolistic coordination, interdependent price elevation is more broadly and effectively deterred. Moreover, as emphasized in the analysis of the paradox of proof in section 3.5, the difference under a direct approach is greatest in industries that are most conducive to substantial price elevation. This beneficial result would, accordingly, make it sensible to pursue a somewhat more relaxed approach toward horizontal mergers, with concomitant gains.⁹⁵

The legal approach toward successful oligopolistic coordination also has implications for competition law rules that relate even more directly to coordinated oligopoly pricing. By analogy to tougher merger enforcement, many likewise advocate an aggressive approach toward facilitating practices (*see, e.g.*, Baker 1993, 207–219).⁹⁶ The same tradeoff is evident: if pricing behavior cannot be reached directly, it becomes necessary to make more practices illegal—whether per se illegal or subject to stricter scrutiny under a rule of reason (stricter because the competitive danger is presumed to be larger). The problem is that there exists, for example, much exchange of information that generates efficiencies, such as in standard-setting and other trade association activities

94 Similar logic also applies to proposals to deconcentrate existing industries.

95 A central thrust of Posner's criticism of Turner, who pushed for a strongly prophylactic merger policy both in his writing and as chief of the Department of Justice Antitrust Division, is that the resulting inefficiency was a further cost of lax treatment of oligopoly pricing. *See* Posner (1969, 1566, 1598–1605); *see also* Demsetz (1973, 3) (suggesting a preference for attacking collusion over tough anti-merger policies that would decrease efficiency).

Relating merger law to price-fixing law raises additional conundrums, at least under a regime like that in the United States which permits private suits for damages. Since a merger itself is clearly a "contract, combination . . . , or conspiracy," the agreement requirement of Sherman Act Section 1 is satisfied, so a demonstration that the merger elevated prices would indicate a violation and hence support a claim for treble damages from the resulting overcharge. Just as mergers may be blocked by the government as anticompetitive if they significantly raise the prospects for coordinated pricing, regardless of the means by which coordination might be accomplished, so pure interdependence would provide a basis for subsequent damages. Accordingly, after mergers, existing law seems in principle to allow private suits for successful oligopolistic coordination, as long as the merger can be shown to contribute to the coordination.

96 *See also* Kaplow (2011c, sec. II.B.2) (further discussing the subject and offering additional references to the literature).

concerned with technological developments, forecasts of demand, and predictions of costs.⁹⁷

4.3. Electronic Communications

Advances in communications technology may alter the detection problem. Indeed, changes in the last decade or two are already changing the landscape.⁹⁸ One problem is that new technologies make rapid, highly detailed price signaling easier. Prices, or elaborate price schemes, can be posted, modified, and rescinded at little cost and in ways that few would notice aside from onlooking competitors (thereby avoiding interim sales transactions). In addition to prices themselves, other product attributes, views about future trends (that may be code for pricing invitations and responses), and other information can be posted on web sites. Firm representatives can participate in (possibly facially anonymous) blogs, chat rooms, and the like. And even old-fashioned press conferences can be replaced with or supplemented by webcasts, greetings, and other online messages.

If all such behavior is to be permitted—that is, not included in the set *X* of prohibited acts—then successful oligopolistic coordination through legally privileged means becomes easier. One implication might be a broader paradox region. On the other hand, if all such behavior is reached, then it is hard to see the limit short of prohibiting all interdependence. It seems difficult to chart a middle course of attempting to delineate a substantial portion of such activity as included in the set *X* while placing the rest in *X'*—and perhaps pointless, as long as enough is in the latter. Furthermore, the speed and unpredictability of future developments may cause newly refined rules tied to particular sorts of communication to be obsolete on arrival.

A second problem concerns detection. To the extent that one does not rely much on inferences from circumstantial evidence—raising the paradox of proof, or being disallowed under the alternative approach sketched in section 4.1—and thus depends primarily on obtaining smoking-gun evidence or near equivalents, the task may become more difficult over time. Detailed codes might be communicated through delays of various numbers of milliseconds in price or other web postings. And the use of VOIP (voice over Internet protocol) or other developing technologies may allow direct communications

97 Note further that, as one expands the set of facilitating practices that are made illegal, which can be seen as broadening the set *X*, the paradox region (e.g., in Figures 2 and 3) shrinks, but, as explained in subsection 3.4.4, once the paradox region becomes narrow, it is hard to understand why one would insist on the acts-based prohibition.

98 Although not drawing the implications considered here, concerns are noted in Baker (1996) and Borenstein (2004).

that are nearly impossible to trace; cell phones already pose significant challenges for law enforcement.

The more public methods discussed just before are also relevant. Currently, many attempts at coordination require multiple meetings over time to account for various complications and changing conditions. But with new technology, it may be possible to have very few initial meetings to choose legally permitted (or illegal but extremely difficult to trace) ways to conduct future interchanges. By greatly reducing in-person and other potentially detectable contact, the likelihood of discovery could fall substantially.

Not all advances in communications, however, make detection more difficult. As email use has become compulsive for many people, individuals engaged in questionable or illegal activity may leave more traces, ones that advanced search algorithms may be able to locate afterwards. For example, in the Microsoft cases in the 1990s, despite presumably sophisticated legal advice, many damaging emails were produced—and ultimately discovered—whereas in the past such communications might have been primarily oral (*see, e.g.,* Heilemann 2000). Likewise with the Whole Foods merger (*see, e.g.,* Levy 2007).

It is unclear which of these features will become more important in the future. If permitted methods of communication become sufficient or if prohibited communications can readily be routed through undetectable channels, the already serious problem of detection may become worse. This concern is especially great if increasingly elaborate legal channels become available, because even finding email trails documenting their use would be to no avail.

Note, in contrast, that direct targeting of successful oligopolistic coordination tends to mitigate these problems. Market-based evidence will continue to be available. Internal evidence may be somewhat harder to come by, but it may still be very useful as long as large numbers of employees (and sometimes consultants or other agents) need to be involved in planning, making, and executing firms' decisions. Indeed, the main force easing detection—the growth of email and other channels that are difficult to erase—may enhance the power of internal evidence in determining whether coordinated oligopoly pricing took place, overcoming efforts by lawyers and others to mask firms' analysis and decision-making. As a result, these technological developments may, as a whole, make a selective communications prohibition even less attractive over time.

4.4. Dispositive Motions

Because so much of the total expense of adjudication is borne before an actual trial commences (investigation, discovery, experts), significant benefits are realized if meritless cases can be dismissed up front. The problem is that,

at the outset, there is very little basis for ascertaining whether cases of the sort examined here have merit. Even under the communications-based prohibition, as long as proof by circumstantial evidence is allowed, an enforcer (private party or government agent) would be in a position to allege that, taking into account all manner of evidence concerning whether the market in question is conducive to successful oligopolistic coordination and whether such successful coordination is taking place, the use of at least some act in X is sufficiently likely under the prevailing legal standard. Without reviewing all of the pertinent evidence (which the enforcer is presumed not to have expended much effort in gathering at the motion-to-dismiss stage), how is an adjudicator to assess such a claim? Moreover, as developed in subsection 3.4.3, under the communications-based prohibition virtually any piece of evidence and thus virtually any factual allegation could favor either a finding of liability or of no liability, depending on the strength with which the allegation is deemed to hold as well as the strength of other allegations and evidence.

This seeming impossibility of rendering a negative judgment at so preliminary a stage is part of what lends appeal to the alternative discussed in section 4.1 of prohibiting all circumstantial evidence—requiring direct evidence of prohibited communications, and possibly even smoking-gun evidence. However, in the absence of an informer who comes forward early in the process,⁹⁹ it is difficult to see how even this narrower question can be resolved at the beginning. After all, we are considering whether a case can be removed from the system before any discovery or extensive investigation by an enforcer. But if cases rise or fall primarily or entirely on the sort of evidence that can be gleaned, not from pricing or other behavior that might be publicly observable, but only from internal evidence deeply hidden in the defendant firms, it would seem that either all cases (except those with an informant) would need to be dismissed, or that essentially none could be disposed of at this preliminary stage.¹⁰⁰ An alternative approach would be to rely more on market-based evidence, some of which is publicly available, at this preliminary stage, switching to the other at

99 Importantly, a significant impetus to informers coming forward is the fear of successful prosecution. But if prosecution is rendered essentially impossible without informants, and if prospective informants know this, informants would be harder to come by.

100 The U.S. Supreme Court in *Twombly* acted as if it had an answer, but what it might be does not seem possible to extract from the opinions in that case (which do not directly consider the dilemma discussed in the text). See Kaplow (2011c, secs. III.B.2, III.E and n.199). This problem is the subject of ongoing research by the author.

later stages (if a communications-based prohibition governs), such as judgment just before trial or in the adjudication itself.¹⁰¹

It remains to consider the disposal of cases after all investigation, discovery, and preparation of expert reports, but before trial. This possibility is especially important when the just-described, more preliminary proceeding cannot screen very powerfully. However, it too is more difficult than is usually imagined. Unless a party's case has negligible evidentiary support, it seems difficult for an adjudicator to find that a trial is unnecessary without in essence conducting the trial to determine whether the evidence is sufficiently strong. In the United States, courts often grant defendants' pre-trial motions for summary judgment in antitrust cases, and a reading of the opinions suggests that some in essence do resolve substantial factual disputes that, in principle, they are supposed to avoid—by denying the motion and allowing trial to proceed. One limitation in achievable cost savings is that such decisions come only after many costs are sunk. Also, because judges are not supposed to resolve cases in this manner, they cannot organize proceedings in a way that best enables them to make the requisite decisions.¹⁰²

4.5. Fit with Legal Doctrine

As stated in the Introduction, this article compares price-fixing regimes in terms of their ability to maximize welfare and does not address existing law, which is the subject of a detailed treatment in a preceding work (Kaplow 2011c), only a few points of which will be mentioned here. This section nevertheless briefly considers the fit of different approaches with legal doctrine.

On one hand, many commentators and recent court decisions in the United States and European Union state that existing law is best understood as something akin to the selective communications prohibition articulated in section 3.1. From the outset, however, there are caveats. To begin, most state the rule instead as one requiring some sort of express or explicit agreement or concert of action, although, as explained, it seems that a selective communications

101 Intelligibly assessing the former at a highly preliminary stage (or even on the eve of trial, as discussed next) seems quite difficult. Yet similar demands are made in other areas of competition law, such as when market power screens are employed. It would seem that, except when an enforcer's claim is transparently weak (which may depend on poor lawyering as well as a lack of factual basis), it would be almost impossible to dispose of a case without some (and possibly significant) assessment of parties' competing factual claims.

102 Sometimes courts also dismiss cases at the outset when factual disputes exist, essentially inferring from the weakness of a plaintiff's complaint that it seems unlikely that its case is substantial. To the extent that this practice becomes more common after *Twombly*, perhaps plaintiffs will begin to offer expert reports and other evidence at this early stage, even though such proffers are supposed to be unnecessary. Preliminary studies of the effects of *Twombly* are described in Kaplow (2011c, 738 n.139).

prohibition is a reasonably close approximation to what is meant.¹⁰³ Second and related, those stating the law in this manner offer neither a reasonably clear operational statement of the rule nor a series of crisp examples from which the contours may readily be inferred. They differ on whether tacit agreements are included, but do not in any case define the concept, and in other respects they employ key terms that are ambiguous or even susceptible to opposite interpretations.

On the other hand, if one considers the actual language of typical prohibitions, which often include terms like conspiracy or concerted action, one is led to a broad prohibition of interdependent oligopolistic behavior. Essentially, these terms focus our attention on the existence of a meeting of the minds, not (contrary to a communications-based prohibition) on the use of particular acts that might be employed to achieve such a mutual understanding. Interestingly, Donald Turner's (1962, 665, 671) seminal article that advances the view that the law's prohibition should be circumscribed as a matter of policy nevertheless endorses this broader interpretation of Sherman Act Section 1's language: "It is not novel conspiracy doctrine to say that agreement can be signified by action as well as by words . . . [T]here are far better grounds for saying that though there may be 'agreement' it is not unlawful agreement."¹⁰⁴ Indeed, the statutes, the general law of conspiracy, and numerous authoritative, earlier Supreme Court decisions support this view. Moreover, the broader interpretation that encompasses interdependent oligopoly behavior is more consistent with the actual practices of lower courts regarding their use of so-called plus factors to find a conspiracy, their instructions to juries on liability, and their method of determining damages. In addition, litigants' behavior with regard to the paradox of proof (as suggested in subsection 3.4.3), as well as

103 See also Kaplow (2011c, sec. III.B.2) and *supra* note 55 (discussing how a communications-based prohibition is akin to treating the proscribed acts as facilitating practices, perhaps eschewing any agreement requirement with respect to their use, while implicitly deeming price fixing itself to be legal).

104 Turner (1962, 683) concludes on the matter as follows:

I also find considerable appeal, as a general matter, in defining "agreement" for purposes of Sherman Act law in terms of interdependence of decisions, if for no other reason than that it seems to me to be a clearer and more workable standard than any other standard, of acceptable scope, which requires something more. Once one goes beyond the boundaries of explicit, verbally communicated assent to a common course of action—a step long since taken and from which it would not seem reasonable to retreat—it is extraordinarily difficult if not impossible to define clearly a plausible limit short of interdependence.

It is ironic, therefore, that the *Twombly* Court, 550 U.S. at 554, cites Turner (1962, 672), for the proposition: "[M]ere interdependence of basic price decisions is not conspiracy."

practitioners' advice and firms' actions, suggests that many hold a broad de facto view of the scope of liability.¹⁰⁵

Finally, the matter needs to be considered in the context of courts' and agencies' overall approach toward competition policy in the recent era, which takes statutory provisions more as invitations to formulate sound rules¹⁰⁶ and, in developing these rules and offering interpretations, reflects an endorsement (stronger in the United States, but growing in the European Union and elsewhere) of economically grounded analysis in lieu of formalistic thinking. As mentioned in the Introduction, the U.S. Supreme Court has strongly and repeatedly advanced this approach, most notably in decisions that overrule precedents.¹⁰⁷ Hence, even if a literal interpretation and prior decisions did suggest otherwise—which is hardly apparent—it is not clear that they would pose a significant obstacle to reform if it came to be understood that views favoring a narrow communications-based prohibition were mistaken. There tends to be a significant lag between theoretical and empirical developments in economic understanding and their dissemination into policy circles and legal advocacy, which eventually translates them into revised interpretations of competition law. With regard to oligopoly pricing, the interval seems longer than usual (despite Posner's early paper and book chapter), a shortcoming that is most appropriately attributable to the behavior of academics, government agencies, and lawyers rather than to the courts.¹⁰⁸

4.6. Comments on Prior Literature

It is surprising that most prior commentary seems to have reached a consensus in favor of a communications-based prohibition without defining it or systematically comparing it to the more direct approach considered here. Indeed, there is very little overlap in content between most prior analysis of how competition policy should address coordinated oligopoly pricing and the present treatment. The most notable exception, as mentioned, is Posner's writing from decades ago, whose existence is frequently acknowledged but whose substance is usually ignored. For the most part, commentators rely on three arguments: it is infeasible to command irrational behavior (abstinence from coordinated price

105 See Kaplow (2011c, part III (also addressing EU Article 101, formerly 81)).

106 For the United States, see, for example, *Leegin*, 551 U.S. at 899, Areeda and Hovenkamp (2003, 226), Baxter (1982), and Kintner (1980, 166, 217, 239).

107 Two of the three cases mentioned there, *Sylvania* and *Leegin*, expressly overruled precedents.

108 For example, with regard to the Supreme Court's *Brooke Group* decision, Baker (1994, 602 and n.84) points out that the plaintiff's submissions and the Court neglected modern economic literature on oligopoly theory that was pertinent to the question at hand.

elevation); a direct prohibition would be inconsistent with the legality of price elevation by monopolists; and injunctive relief would be tantamount to price regulation. Even in extensive treatments of competition law toward price fixing, these points are usually asserted in a few sentences or at most a few paragraphs.¹⁰⁹ Although it has been explained why the first two arguments are confused and the third has been implicitly addressed at length, it is worth confronting each directly here.

The argument that a prohibition on successful oligopolistic coordination—without a further requirement that particular acts were employed—would entail commanding firms to behave irrationally is addressed in section 2.3. The simple response is deterrence: the threat of sanctions changes what firms find to be in their interest to do (as long as the expected punishment exceeds the profits from oligopolistic price elevation). In this respect, coordinated pricing is little different from unsafe disposal of toxic waste or cheating on taxes, activities that profit-maximizing firms find it in their rational interest to undertake, that is, unless there are substantial penalties.

The second argument is that prohibiting oligopolistic coordination and thus the resulting price elevation is inconsistent with permitting monopolists to charge supracompetitive prices. Subsection 2.1.2 explains the rationale for differential treatment. Indeed, the justification for permitting firms to reap the rewards of dynamically efficient behavior while denying them profits from higher prices attributable merely to their refraining from competition is entirely familiar. Moreover, the logic of this objection, as noted before, applies equally to classic cartels. That is, if oligopolists should be free to raise prices because monopolists are permitted to do so, the logic applies not only to purely interdependent coordination but also to that facilitated by additional means of communication such as meetings in hotel rooms. Of course, none who offer this argument favor making cartels legal.

The third objection is that remedies, with a focus on injunctive relief, would entail all the difficulties associated with price regulation.¹¹⁰ As just mentioned, this claim is confused because it overlooks deterrence through the prospect of

109 See sources cited *supra* notes 23 and 46 and *infra* note 110.

110 See, e.g., Turner (1962, 669) (referring to a “purely public-utility interpretation of the Sherman Act”); Areeda & Hovenkamp (2003, 206) (noting “the absence of a practical remedy other than judicial price control”); *id.* at 232–233 (suggesting that the required injunction would be “equivalent to . . . compelling marginal-cost pricing” which “puts the antitrust tribunal directly in the price-control business”); *id.* at 273 (describing the FTC’s view in *du Pont*); Baker (1996, 47) (“the only remedy is judicial price regulation—a complete non-starter”); Elhauge & Geradin (2007, 835) (“Could courts figure out whether they had done so other than by asking whether the prices were reasonable?”); and Scherer (1977, 984).

penalties.¹¹¹ Also, as explained in section 2.3, it is unclear that injunctions have an important role to play in attacking successful oligopolistic coordination. Instead, the real question involves the feasibility of detection: if detection is highly imperfect and involves a substantial risk of false positives, socially costly chilling effects will ensue.

Of course, the concern about detection and chilling effects—and related matters of the appropriate burden of proof and the use of varying types of evidence of differing reliability—has consumed much of the attention in this article, beginning with subsection 2.1.3, developed in section 2.2, and applied to a communications-based prohibition in sections 3.3 through 3.5. Among the highlights are: chilling effects indeed constitute one of the two central concerns (along with deterrence) in designing a legal regime addressed to coordinated oligopoly pricing; as a consequence, it is important to consider the full range of types of evidence, employing appropriate mixes in particular cases rather than arbitrarily confining attention to a subset that may or may not be reliable in a given application; and evidentiary standards should depend on the nature of proof. Moreover, systematic comparison in section 3.5 with commentators' preferred communications-based prohibition is not favorable to it in this regard, which nearly everyone seems to take for granted: defendants are exonerated in situations in which the social danger is greatest and chilling costs are likely to be low, whereas they may be found liable when the danger is modest and the risk of chilling desirable behavior is high—the reverse of what would be optimal. Additionally, a communications-based prohibition may well be more difficult to implement, perhaps substantially so. These matters are not examined in prior treatments of the subject that advance the objection under consideration.

5. CONCLUSION

Determination of an optimal regime to combat coordinated oligopolistic price elevation requires attention to the social objective, the problem of detection, and the design of sanctions. The social purpose of the regime is to deter such price elevation. Successful oligopolistic coordination may be detected through a number of means, ideally considered in combination. Its presence or absence may be demonstrated by market-based evidence, notably pricing patterns and indications of price elevation, and also by the use of facilitating practices. The degree to which industry conditions are conducive to collusion is also highly

111 Furthermore, as discussed in section 3.6, penalties are also required under a communications-based prohibition, and fines (and/or private damage awards) really should be geared to the magnitude of oligopolists' price elevation, so the need for such measurement is not avoided, only deferred (and limited to cases that happen to involve demonstrable prohibited acts).

relevant. In a sense, conduciveness is a necessary condition, but not a sufficient one—because of the possibility that deterrence is effective, among other reasons. Interfirm communications, the emphasis of much prior analysis of the subject, can be probative but are not necessary, both as a logical matter and because proof may be lacking even when they are present. All bases for assessment may be established, reinforced, or negated by evidence internal to firms, including many sorts that do not speak directly to coordination. For example, if the main alternative explanation for a sudden price increase is a cost shock, firms' internal records may well indicate whether such an event occurred.

The analysis of sanctions focuses on fines and private damage awards because monetary sanctions are heavily used and tend to be the most efficient tool. Imprisonment can provide a helpful supplement in achieving deterrence. Injunctions do not seem to be widely used and do not appear to be important in principle because they fail to generate deterrence and must be enforced through other sanctions in any event. Nevertheless, many writers seem to be preoccupied with injunctions and believe that rules of liability need to be crafted, and possibly significantly constrained, by the feasibility of injunctions of a particular type, tantamount to command-and-control regulation.

Most commentators—legal analysts and economists, in the European Union and the United States—believe that current law is best described as and should be a communications-based prohibition, interpreting notions of agreement, conspiracy, and concerted action so as to limit liability to cases in which the use of particular, explicit forms of communication have been employed. Operational statements of the content of the imagined rule are difficult to extract, and there are serious questions regarding the aptness of their characterization of existing law, both subjects of a predecessor to this article (Kaplow 2011c). Perhaps most notable, conventional policy discussion of this subject has little overlap with the analysis offered here, which should raise serious doubt about whether the dominant view is well grounded. Systematic examination of the social objective, the problem of detection, and the matter of appropriate sanctions does not generate the commonly favored rule, indicating that it indeed is not optimal. Nevertheless, this article is devoted to explicit comparison of the direct approach with the widely endorsed alternative.

The communications-based prohibition considers only a subset of means to a socially undesirable end rather than focusing on the end itself, deterrence, and on other matters of direct concern, notably, chilling effects (although the latter obviously motivate those advancing the approach). Its main deficiencies can be traced to this core trait. To begin, the communications-based prohibition is defined in a formalistic way that is significantly removed from the social objective. Indeed, this is an inherent feature, for under a functional approach it dissolves into the direct method, which its proponents reject.

Regarding the central challenge of detection, the communications-based prohibition utilizes a large body of relevant types of evidence only indirectly, even when it is highly probative of the ultimate social question, and considers another portion of evidence (on conduciveness of conditions) to bear on liability in a manner opposite to its implications for social welfare. Hence, even before undertaking a detailed inquiry, the dominant view appears dubious. When one focuses on the wrong question, one is unlikely to obtain a good answer, much less the best one—and this logic holds even when the correct answer is difficult to ascertain due to limited empirical knowledge and the complexity of the task. To be sure, detection is not particularly problematic in those cases that generate smoking-gun evidence, such as tapes of secret price-fixing meetings, but such evidence would likewise be credited under the more direct approach to liability. Serious defects, however, become apparent when considering proof by circumstantial evidence, which is widely allowed and thought to be essential in light of firms' efforts to hide their actions.

As explained at length, the logical inference process entailed by a communications-based prohibition gives rise to a paradox of proof that has startling implications that do not seem well appreciated. The proof process is highly complex and depends on knowledge that does not exist and is unlikely to become available. It implies that enforcers and defendants would routinely adopt strategies in adjudication that are not widely employed. Most important, in comparing the two approaches, the communications-based prohibition is relatively more likely to find liability when deterrence benefits are low and chilling costs are high, while it exonerates defendants when deterrence benefits are greatest and chilling costs are low. Accordingly, it seems quite unlikely to be an optimal compromise or a sensible proxy criterion for determining when sanctions should be applied. It is explained how the direct method—with a burden of proof calibrated to provide a similar rate of applicability to potential cases—dominates the communications-based prohibition in the sense that the former finds liability in cases characterized both by relatively larger deterrence benefits and relatively smaller chilling costs than under the latter.

There exist some further conundrums. If certain common empirical conjectures are granted, the actual outcomes under a communications-based prohibition would hardly differ from those under the more direct approach, which makes all the more curious the intensity with which proponents insist that it be rejected. Relatedly, if cases with successful oligopolistic coordination and no explicit communications are rare, as many assert,¹¹² then even under a narrow, communications-based prohibition, a private or government enforcer should

112 See *supra* subsection 3.4.4.

prevail (not just at preliminary stages, but even at trial) upon demonstration of successful coordinated price elevation, absent strong proof from defendants that their case is truly exceptional—a point that commentators and courts have overlooked. Additionally, there are serious reasons to question the standard view about the implications of highly conducive conditions, and it is possible that the central difference in the two approaches vanishes. The communications-based prohibition also raises complications in assessing the appropriate magnitude of sanctions. Finally, such a prohibition may be more (perhaps much more) complex and costly to implement than a direct one.

This article also analyzes a doctrinally radical alternative of disallowing all circumstantial evidence. It has a *prima facie* plausibility that the commonly advanced version lacks, although it is difficult to define and is subject to various shortcomings. If it happened to be true that smoking-gun evidence would quickly turn up whenever successful oligopolistic coordination took place, then a highly circumscribed approach may be adequate, but the direct approach would produce this result under such an assumption. However, it is widely believed that this detection strategy by itself is insufficient—and, indeed, it might miss those cases involving the greatest expected social harm. If so, a wider range of detection methods needs to be permitted, and in any event the choice of methods and setting of proof burdens should grow out of a decision-making framework aimed at maximizing social welfare, not at meeting a formalistic requirement.

Brief consideration is also given to the fit of various approaches with prevailing doctrine. Competition authorities and adjudicators, especially the U.S. Supreme Court, have increasingly embraced rules grounded in economic understandings while rejecting formalistic alternatives. Accordingly, if the systematic application of economic analysis to the most important competition rule suggests a need for modification, one should expect the relevant legal authorities to be receptive—and the language and somewhat amorphous interpretation of existing rules may pose little obstacle in any event (Kaplow 2011c). The article also examines the implications of the communications-based prohibition for other competition rules and the impact of ongoing evolution in electronic communications. Finally, the leading arguments in prior literature in favor of a narrow prohibition—each more a brief observation than a developed construct—are related to the present analysis.

This investigation is preliminary in a number of respects. First, there has been little prior analysis of most of the relevant territory: concerning elaboration of the social objective, including the development of chilling costs and the framework for trading off errors; assessment of the many avenues of detection, with further attention to chilling costs and an emphasis on internal as well as other forms of evidence; and the choice among sanctions and setting of their

magnitude. (For further exploration, see Kaplow (2011a).) Second, a number of key decisions, at least as a matter of degree, raise empirical questions about which evidence is limited. Third, the interplay with enforcement institutions needs further attention. It may be, for example, that the optimal legal rule for a system with exclusive government enforcement using an expert agency differs from that for a regime allowing private suits adjudicated by generalist judges and lay juries.

Despite these significant reservations, it does appear that wholesale rethinking is in order. The predecessor to this article (Kaplow 2011c) explains how existing formulations of the legal prohibition are incoherent to a substantial degree and differ greatly from what many advancing them seem to contemplate. Furthermore, consensus characterizations of existing practice are difficult to state operationally and conflict with much of what actually seems to take place.

More relevant for policy purposes, this article offers a systematic comparison of commentators' favored method and a more direct approach. It turns out that the investigation differs even in basic subject matter coverage (not just in possible bottom lines) from most prior work on the topic. When the commonly supported communications-based prohibition is compared to the more direct approach outlined here, sharp differences appear, and in important respects the former seems to have much backwards.

These articles therefore constitute a strong critique of present thinking about competition law directed toward coordinated price elevation. Nevertheless, it should be noted that the criticism herein is not primarily aimed at courts because academics and other analysts are primarily responsible for assessing policy. This article seeks to rekindle and advance a long-dormant policy debate on perhaps the most important subject of competition regulation. It also suggests directions for empirical research that would valuably illuminate the subject and thus better guide future competition policy in this realm.¹¹³

REFERENCES

- ABA Section of Antitrust Law. 2007. *Antitrust Law Developments*, vol. 1, 6th edn. Chicago: ABA Publishing.
- Areeda, Phillip E., & Herbert Hovenkamp. 2003. *Antitrust Law*, vol. 6, 2d edn. New York: Aspen Law & Business.

113 As suggested in section 3.4, if a communications-based prohibition were to be taken seriously, a substantial additional empirical research program would be necessary—to determine the height and shape of the curves depicted there—one that seems more difficult than even the challenging one necessary to address the questions identified in part 2.

- Areeda, Phillip E., Herbert Hovenkamp, Roger D. Blair, & Christine Plette Durrance. 2007. *Antitrust Law*, vol. 2A, 3d edn. Austin: Wolters Kluwer Law & Business.
- Areeda, Phillip, & Louis Kaplow. 1988. *Antitrust Analysis*, 4th edn. Boston: Little, Brown & Company.
- Bagwell, Kyle, & Asher Wolinsky. 2002. Game Theory and Industrial Organization. In Robert J. Aumann & Sergiu Hart, eds., *Handbook of Game Theory with Economic Applications*, vol. 3. Amsterdam: North-Holland.
- Baker, Jonathan B. 1989. Identifying Cartel Policing Under Uncertainty: The U.S. Steel Industry, 1933–1939. 32 *Journal of Law & Economics* S47–S76.
- . 1993. Two Sherman Act Section 1 Dilemmas: Parallel Pricing, the Oligopoly Problem, and Contemporary Economic Theory. 38 *Antitrust Bulletin* 143–219.
- . 1994. Predatory Pricing after Brooke Group: An Economic Perspective. 62 *Antitrust Law Journal* 585–603.
- . 1996. Identifying Horizontal Price Fixing in the Electronic Marketplace. 65 *Antitrust Law Journal* 41–55.
- Baker, Jonathan B., & Timothy F. Bresnahan. 1992. Empirical Methods of Identifying and Measuring Market Power. 61 *Antitrust Law Journal* 3–16.
- . 2008. Economic Evidence in Antitrust: Defining Markets and Measuring Market Power. In Paolo Buccirossi, ed., *Handbook of Antitrust Economics*. Cambridge, MA: MIT Press.
- Baxter, William F. 1982. Separation of Powers, Prosecutorial Discretion, and the “Common Law” Nature of Antitrust Law. 60 *Texas Law Review* 661–703.
- Blair, Roger D. 2008. Introduction. 53 *Antitrust Bulletin* 1–4.
- Bishop, William. 1983. Oligopoly Pricing: A Proposal. 28 *Antitrust Bulletin* 311–336.
- Black, Oliver. 2005. Communication, Concerted Practices and the Oligopoly Problem. 1 *European Competition Journal* 341–346.
- Borenstein, Severin. 2004. Rapid Price Communication and Coordination: The Airline Tariff Publishing Case (1994). In John E. Kwoka, Jr. & Lawrence J. White, eds., *The Antitrust Revolution*, 4th edn. New York: Oxford University Press.
- Bork, Robert H. 1978. *The Antitrust Paradox: A Policy at War with Itself*. New York: Basic Books.
- Bresnahan, Timothy F. 1987. Competition and Collusion in the American Automobile Industry: The 1955 Price War. 35 *Journal of Industrial Economics* 457–482.

- . 1989. Empirical Studies of Industries with Market Power. In Richard Schmalensee & Robert D. Willig, eds., *Handbook of Industrial Organization*, vol. 2. Amsterdam: North-Holland.
- . 1997. Testing and Measurement in Competition Models. In David M. Kreps & Kenneth F. Wallis, eds., *Advances in Economics and Econometrics: Theory and Applications, Seventh World Congress*, vol. 3. Cambridge, UK: Cambridge University Press.
- Brock, James W. 2006. Antitrust Policy and the Oligopoly Problem, 51 *Antitrust Bulletin* 227–280.
- Buccirossi, Paolo. 2008. Facilitating Practices. In Paolo Buccirossi, ed., *Handbook of Antitrust Economics*. Cambridge, MA: MIT Press.
- Chamberlin, Edward H. 1929. Duopoly: Value Where Sellers Are Few. 44 *Quarterly Journal of Economics* 63–100.
- Comanor, William S., & Mark A. Schankerman. 1976. Identical Bids and Cartel Behavior. 7 *Bell Journal of Economics* 281–286.
- Connor, John M. 2007. Price-Fixing Overcharges: Legal and Economic Evidence. 22 *Research in Law & Economics* 59–154.
- . 2008. *Global Price Fixing*, 2d updated & rev. edn. Berlin: Springer-Verlag.
- Cook, Paul W., Jr. 1963. Fact and Fancy on Identical Bids. 41 (Jan.–Feb.) *Harvard Business Review* 67–72.
- Crandall, Robert W., & Clifford Winston. 2003. Does Antitrust Policy Improve Consumer Welfare? Assessing the Evidence. 17 (Fall) *Journal of Economic Perspectives* 3–26.
- Dabbah, Maher M. 2004. *EC and UK Competition Law*. Cambridge, UK: Cambridge University Press.
- Davis, Douglas D., & Bart J. Wilson. 2002. Experimental Methods and Antitrust Policy. In Charles A. Holt & R. Mark Isaac, eds., *Experiments Investigating Market Power*. New York: Elsevier Science JAI.
- Demsetz, Harold. 1973. Industry Structure, Market Rivalry, and Public Policy. 16 *Journal of Law & Economics* 1–9.
- Department of Justice, Antitrust Division. 2010. *Price Fixing, Bid Rigging, and Market Allocation Schemes: What They Are and What to Look For*. Available at: <http://www.justice.gov/atr/public/guidelines/211578.pdf>.
- Easterbrook, Frank. 1984. The Limits of Antitrust. 63 *Texas Law Review* 1–40.
- Elhauge, Einer, & Damien Geradin. 2007. *Global Antitrust Law and Economics*. New York: Foundation Press.
- Elzinga, Kenneth G. 1984. New Developments on the Cartel Front. 29 *Antitrust Bulletin* 3–26.
- Fraas, Arthur G., & Douglas F. Greer. 1977. Market Structure and Price Collusion: An Empirical Analysis. 26 *Journal of Industrial Economics* 21–44.

- Friedman, James W. 1971. A Non-cooperative Equilibrium for Supergames. 38 *Review of Economic Studies* 1–12.
- . 1986. *Game Theory with Applications to Economics*. New York: Oxford University Press.
- Fudenberg, Drew, & Jean Tirole. 1991. *Game Theory*. Cambridge, MA: MIT Press.
- Green, Edward J., & Robert H. Porter. 1984. Noncooperative Collusion under Imperfect Price Information. 52 *Econometrica* 87–103.
- Hall, Robert E. 2007. (Review of Whinston, *Lectures on Antitrust Economics*.) 45 *Journal of Economic Literature* 1066–1070.
- Harrington, Joseph E., Jr. 2006. How Do Cartels Operate? 2 *Foundations & Trends in Microeconomics* 1–105.
- . 2008. Detecting Cartels. In Paolo Buccirossi, ed., *Handbook of Antitrust Economics*. Cambridge, MA: MIT Press.
- Hay, George A. 1981. Oligopoly, Shared Monopoly, and Antitrust Law. 67 *Cornell Law Review* 439–481.
- . 2006. Horizontal Agreements: Concept and Proof. 51 *Antitrust Bulletin* 877–914.
- Hay, George A., & Daniel Kelley. 1974. An Empirical Survey of Price-Fixing Conspiracies. 17 *Journal of Law & Economics* 13–38.
- Heilemann, John. 2000. The Truth, The Whole Truth, and Nothing But The Truth: The Untold Story of the Microsoft Antitrust Case and What It Means for the Future of Bill Gates and His Company, *Wired*, November. Available at: http://www.wired.com/wired/archive/8.11/microsoft_pr.html.
- Hovenkamp, Herbert. 2005. *The Antitrust Enterprise*. Cambridge, MA: Harvard University Press.
- Jacquemin, Alexis, & Margaret E. Slade. 1989. Cartels, Collusion, and Horizontal Merger. In Richard Schmalensee & Robert D. Willig, eds., *Handbook of Industrial Organization*, vol. 1. Amsterdam: North-Holland.
- Jones, Alison. 1993. Woodpulp: Concerted Practice and/or Conscious Parallelism? 14 *European Competition Law Review* 273–279.
- Kaplow, Louis. 2011a. An Economic Approach to Price Fixing. 77 *Antitrust Law Journal* 343–449.
- . 2011b. On the Choice of Welfare Standards in Competition Law. In Daniel Zimmer, ed., *Goals of Competition Law*. Cheltenham: Edward Elgar (forthcoming).
- . 2011c. On the Meaning of Horizontal Agreements in Competition Law. 99 *California Law Review* 683–818.
- . 2012a. Burden of Proof. 121 *Yale Law Journal* 738–859.
- . 2012b. On the Optimal Burden of Proof. *Journal of Political Economy* (forthcoming).

- Kaplow, Louis, & Carl Shapiro. 2007. Antitrust. In A. Mitchell Polinsky & Steven Shavell, eds., *Handbook of Law and Economics*, vol. 2. Amsterdam: Elsevier.
- Kaplow, Louis, & Steven Shavell. 1994. Why the Legal System Is Less Efficient than the Income Tax in Redistributing Income. 23 *Journal of Legal Studies* 667–681.
- Kaysen, Carl. 1951. Collusion under the Sherman Act. 65 *Quarterly Journal of Economics* 263–270.
- Kaysen, Carl, & Donald F. Turner. 1959. *Antitrust Policy: An Economic and Legal Analysis*. Cambridge, MA: Harvard University Press.
- Kintner, Earl. 1980. *Federal Antitrust Law*. Cincinnati: Anderson Publishing Company.
- Kovacic, William E. 1993. The Identification and Proof of Horizontal Agreements under the Antitrust Laws. 38 *Antitrust Bulletin* 5–81.
- Kühn, Kai-Uwe. 2001. Fighting Collusion by Regulating Communication between Firms. 16 *Economic Policy* 167–204.
- Kühn, Kai-Uwe, & Xavier Vives. 1995. *Information Exchanges Among Firms and their Impact on Competition*. Luxembourg: European Community.
- LaFave, Wayne R. 2003. *Substantive Criminal Law*, vol. 2, 2d edn. Eagan, MN: Thompson/West.
- Levenstein, Margaret C. 1997. Price Wars and the Stability of Collusion: A Study of the Pre-World War I Bromine Industry. 45 *Journal of Industrial Economics* 117–138.
- Levy, Judith. 2007. FTC: Whole Foods–Wild Oats Merger Would Thwart Competition, *Seeking Alpha*, June 20. Available at: <http://seekingalpha.com/article/38876-ftc-whole-foods-wild-oats-merger-would-thwart-competition>.
- Lopatka, John E. 1996. Solving the Oligopoly Problem: Turner's Try. 41 *Antitrust Bulletin* 843–908.
- Mankiw, N. Gregory, & Michael D. Whinston. 1986. Free Entry and Social Inefficiency. 17 *RAND Journal of Economics* 48–58.
- Markovits, Richard S. 1974. Oligopolistic Pricing Suits, the Sherman Act, and Economic Welfare: Part I: Oligopolistic Price and Oligopolistic Pricing: Their Conventional and Operational Definition. 26 *Stanford Law Review* 493–548.
- . 1975. Oligopolistic Pricing Suits, the Sherman Act, and Economic Welfare: Part III: Proving (Illegal) Oligopolistic Pricing: A Description of the Necessary Evidence and a Critique of the Received Wisdom About Its Character and Cost. 27 *Stanford Law Review* 307–332.
- Masson, Robert T., & Robert J. Reynolds. 1978. Statistical Studies of Antitrust Enforcement: A Critique. In American Statistical Association, 1977 *Proceedings of the Business and Economics Statistics Section*, part 1.

- McAfee, R. Preston, & John McMillan. 1992. Bidding Rings. 82 *American Economic Review* 579–599.
- Monti, Giorgio. 2001. The Scope of Collective Dominance under Articles 82 EC. 38 *Common Market Law Review* 131–157.
- Motta, Massimo. 2004. *Competition Policy*. Cambridge, UK: Cambridge University Press.
- . 2007. Review of Michael Whinston, *Lectures on Antitrust Economics*. 3 *Competition Policy International* 313–320.
- Neven, Damien J. 2001. “Collusion” under Article 81 and the Merger Regulation. In Konkurrensverket, Swedish Competition Authority, *Fighting Cartels – Why and How?*. Göteborg: Elanders Graphic Systems.
- Nye, Stephen A. 1975. Can Conduct Oriented Enforcement Inhibit Conscious Parallelism? 44 *Antitrust Law Journal* 206–230.
- Osborne, Martin J., & Ariel Rubinstein. 1994. *A Course in Game Theory*. Cambridge, MA: MIT Press.
- Perloff, Jeffrey M., Larry S. Karp, & Amos Golan. 2007. *Estimating Market Power and Strategies*. Cambridge, UK: Cambridge University Press.
- Polinsky, A. Mitchell, & Steven Shavell. 2007. The Theory of Public Enforcement of Law. In A. Mitchell Polinsky & Steven Shavell, eds., *Handbook of Law and Economics*, vol. 1. Amsterdam: Elsevier.
- Porter, Robert H. 1983a. Optimal Cartel Trigger Price Strategies. 29 *Journal of Economic Theory* 313–338.
- . 1983b. A Study of Cartel Stability: The Joint Executive Committee. 14 *Bell Journal of Economics* 301–314.
- . 2005. Detecting Collusion. 26 *Review of Industrial Organization* 147–167.
- Porter, Robert H., & J. Douglas Zona. 1993. Detection of Bid Rigging in Procurement Auctions. 101 *Journal of Political Economy* 518–539.
- . 2008. Collusion. In Wayne Dale Collins, ed., *Issues in Competition Law and Policy*, vol. 2. Chicago: ABA Publishing.
- Posner, Richard A. 1969. Oligopoly and the Antitrust Laws: A Suggested Approach. 21 *Stanford Law Review* 1562–1606.
- . 1976. Oligopolistic Pricing Suits, the Sherman Act, and Economic Welfare: A Reply to Professor Markovits. 28 *Stanford Law Review* 903–914.
- . 2001. *Antitrust Law*, 2d edn. Chicago: University of Chicago Press.
- Posner, Richard A., & Frank H. Easterbrook. 1981. *Antitrust: Cases, Economic Notes and Other Materials*, 2d edn. Eagan, MN: West Publishing Co.
- Sagi, Guy. 2008. The Oligopolistic Pricing Problem: A Suggested Price Freeze Remedy. 2008 *Columbia Business Law Review* 269–359.
- Scherer, F.M. 1977. The Posnerian Harvest: Separating Wheat from Chaff. 86 *Yale Law Journal* 974–1002.

- Scherer, F.M., & David Ross. 1990. *Industrial Market Structure and Economic Performance*, 3d edn. Boston: Houghton Mifflin Co.
- Schmalensee, Richard. 1987a. Competitive Advantage and Collusive Optima. 5 *International Journal of Industrial Organization* 351–367.
- . 1987b. Collusion versus Differential Efficiency: Testing Alternative Hypotheses. 35 *Journal of Industrial Economics* 399–426.
- . 1989. Inter-Industry Studies of Structure and Performance. In Richard Schmalensee & Robert D. Willig, eds., *Handbook of Industrial Organization*, vol. 2. Amsterdam: North-Holland.
- Shapiro, Carl. 1989. Theories of Oligopoly Behavior. In Richard Schmalensee & Robert D. Willig, eds., *Handbook of Industrial Organization*, vol. 1. Amsterdam: North-Holland.
- Slade, Margaret E. 1990. Strategic Pricing Models and Interpretation of Price-War Data. 34 *European Economic Review* 524–537.
- Spence, A. Michael. 1978. Tacit Co-ordination and Imperfect Information. 3 *Canadian Journal of Economics* 490–505.
- Stigler, George J. 1964. A Theory of Oligopoly. 72 *Journal of Political Economy* 44–61.
- Stroux, Sigrid. 2004. *US and EC Oligopoly Control*. The Hague: Kluwer Law International.
- Tirole, Jean. 1988. *The Theory of Industrial Organization*. Cambridge, MA: MIT Press.
- Turner, Donald F. 1962. The Definition of Agreement under the Sherman Act: Conscious Parallelism and Refusals to Deal. 75 *Harvard Law Review* 655–706.
- Van Bael, Ivo, & Jean-Francois Bellis. 2005. *Competition Law of the European Community*, 4th edn. The Hague: Kluwer.
- Vives, Xavier. 1999. *Oligopoly Pricing: Old Ideas and New Tools*. Cambridge, MA: MIT Press.
- Werden, Gregory J. 2004. Economic Evidence on the Existence of Collusion: Reconciling Antitrust Law with Oligopoly Theory. 71 *Antitrust Law Journal* 719–800.
- Whinston, Michael D. 2006. *Lectures on Antitrust Economics*. Cambridge, MA: MIT Press.
- White House Task Force on Antitrust Policy. 1968 (reprinted in *Antitrust & Trade Regulation Reporter*, no. 411, spec. supp. II, May 27, 1969).
- White, Lawrence J. 2001. Lysine and Price Fixing: How Long? How Severe? 18 *Review of Industrial Organization* 23–32.